DOCUMENT RESUME

ED 038 463 UD 009 939

AUTHOR	Schumer, Florence; Deutsch, Cynthia P.
TITLE	An Exploration of Transfer of Independent Behavior from Enrichment Classrooms in Harlem to a Non-Classroom Behavioral Situation. Final Report for Interim Research Period, September 1, 1968 to August 31, 1969.
INSTITUTION	New York Univ., N.Y. Inst. for Developmental Studies.
SPONS AGENCY	Office of Economic Opportunity, Washington, D.C.
PUB DATE	Dec 69
NOTE	99p.
EDRS PRICE	EDRS Price MF-$0.50 HC Not Available from EDRS.
DESCRIPTORS	Check Lists, Classroom Environment, *Classroom Observation Techniques, Conformity, Curriculum Evaluation, Disadvantaged Youth, Educational Objectives, Elementary School Students, *Enrichment Programs, *Individualism, Interaction Process Analysis, Negro Students, Public Schools, *Student Behavior, *Transfer of Training, Urban Youth
IDENTIFIERS	New York City

ABSTRACT

This study concerns a demonstration enrichment program for intervening into the ways of thinking, learning, and behaving displayed by urban ghetto children in Harlem public schools. The study focuses on the implications of ongoing research by the Institute for Developmental Studies (IDS) for the further understanding of the independent behavior of children in and out of the classroom, and also on the effectiveness of the IDS classroom procedures. In the evaluation procedure, behavioral rather than standardized test criteria were used, both in the classroom and in a "transfer" situation; classroom observation followed by rating of third-grade children from the demonstration program on "systematically developed, reliable, behavioral checklists" was used. The findings of the study indicated that the expectations of reliability and usability of the classroom behavior checklist were realized. However, it was found that independence of behavior in children--at least in the age group of the subjects tested--does not "carry over" into situations outside the classroom. Further, the nature of the task, situation, and the requirement placed before the child were found to seem to be of key significance. [Not available in hard copy due to marginal legibility of original document.] (RJ)

ED038463

FINAL REPORT FOR INTERIM RESEARCH PERIOD

September 1st, 1968 to August 31st, 1969

OEO Contract No. B89-4612A

An Exploration of Transfer of Independent Behavior from Enrichment Classrooms in Harlem to a Non-Classroom Behavioral Situation

Florence Schumer, Ph. D. Principal Investigator

Cynthia P. Deutsch, Ph. D. Co-Principal Investigator

Institute for Developmental Studies
School of Education, New York University
Washington Square, New York, N. Y. 10003

Martin Deutsch, Ph. D., Director

UD009939E

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

December, 1969

Project Staff (1968-1969)

Research Coordinator: Marguerite Levy, Ph.D.

Research Assistants:
Jo Renée Fine, M.A.
Diana Lackowitz, M.A.
Peter Pardine, B.A.
Dorothy Schmidt, M.A.
William Swinburne, B.A.
Patricia Wagner, M.A.

Fine, Jo Renée B.A., Smith College, 1965; M.A., Center for Human Relations, New York University, 1967; Doctoral candidate in Education, New York University. Relevant experience includes educational research with specific experience in developing interview schedules for, and interviewing of, ghetto populations.

Lackowitz, Diana B.A., Harpur College, 1966; M.A., University of Connecticut, 1967; Special Student in Educational Psychology, New York University. Relevant experience includes administration of individual intelligence tests to ghetto populations and educational research.

Levy, Marguerite B.A., University of Buffalo, 1952; M.A., New York University, 1964; Ph.D. in Psychology, New York University, 1968. Relevant experience includes teaching and research in general, industrial, and social psychology.

Pardine, Peter B.A., Fairleigh Dickinson University, 1963; Doctoral candidate in Psychology, New York University. Relevant experience includes teaching at the elementary level, as well as research in experimental psychology.

Schmidt, Dorothy B.A., Barnard College, 1958; M.A. in Psychology, Boston University, 1962. Relevant experience includes research in learning psychology; five years of experience at the Institute for Developmental Studies.

Swinburne, William B.A., University of Vermont, 1966; Masters Candidate in Educational Psychology, New York University. Relevant experience includes administration of individual intelligence tests to ghetto populations. Former director of East Harlem Induction Center for Addiction Service Agencies, City of New York.

Wagner, Patricia B.A., Fairleigh Dickinson University, 1966; M.A. in Psychology, Columbia University Teachers College, 1968. Relevant experience includes graduate work in developmental psychology; teaching retarded children; and internship in school psychology at Willowbrook State School.

Acknowledgements

We would like to note that the current report is best regarded as an interim report rather than as a final report, in the light of the fact that our year's work is being continued and expanded. Theoretical integration of the research findings as well as extensive review of relevant literature are neither possible nor appropriate at this point.

We would like very much to thank the children and teachers who have helped us in the execution of our research, and specifically wish to thank the staff and general personnel in the four Harlem schools in which so much of our pilot and experimental activity took place. These schools are: Public Schools 90 on West 148th Street, 175 on West 134th Street, and 79 on East 120th Street. Some of our pilot-testing was also done at the Friends Seminary through the kind permission of its Lower School Principal, Mrs. Joseph Church.

The following people deserve special mention: Mr. Martin Frey and Mr. Edwin Haas, District Superintendents; Dr. Abraham Cohen, Mr. Sol Derewetsky, and Mr. Leroy Watkins, Principals; Mr. George Carter, and Mr. Herbert Coleman, Assistant Principals; Mrs. Frances Alston, Miss Elna Haynes, Mrs. Odessa Jenkins, and Miss Bonnie Rukin, Teachers; and Mrs. Annette Jackson, Miss Monica Raymond, Miss Phyllis Sheldon, and Mr. Robert Weinheimer, Assistant Teachers.

We also wish to express our appreciation to members of our own staff for their help in arranging our entry into the schools: Mrs. Jacqueline Stuchin, Curriculum Director, Miss Edwina Meyers, and Mrs. Laura Schneider Ellis, Supervisors. Dr. Carole Silfen, Research Scientist on our staff, also assisted us in connection with some of the behavioral ratings. Finally,

special thanks are due to Mrs. Elsie Santiago and Mrs. Lynda Hurwitz for their dedicated secretarial assistance throughout the year.

Contents

Page

Chapter 1 Background and Hypotheses of the Current Investigation...... 1

Chapter 2 Sample and Methodology.................................. 9

Chapter 3 Findings.. 31

Chapter 4 Summary and Conclusions; Abstract of Continuation Research.. 45

References 52

Tables 54

Figures 75

Appendix A The Classroom Behavior Checklist.......................... 87

Appendix B The Revised Classroom Behavior Checklist.................. 89

List of Tables

Table

1 Experimental Sample Observed in the Classrooms: Mean Age and Sex for Subjects in Public Schools 90 and 175 Classified According to Years of Exposure to the Enrichment Program

2 Experimental Sample Observed in the Classrooms: Mean Age for Subjects Classified According to Years of Exposure to the Enrichment Program for Both Schools Combined

3 Behavioral Transfer Sessions--Experimental Subjects: Mean Age and Sex for Subjects in Public Schools 90 and 175 Classified According to Years of Exposure to the Enrichment Program

4 Experimental Sample in the Behavioral Transfer Sessions: Mean Age for Subjects Classified According to Years of Exposure to the Enrichment Program for Both Schools Combined

5 Behavioral Transfer Sessions--Control Subjects: Mean Age and Sex for Subjects in Public Schools 90 and 175

6 Distribution of Subjects by School and Sex in the Transfer Sessions

7 Inter-rater Reliability Coefficients of Behavior Checklist Items

8 Frequencies of No Opportunity to Observe Category for Behavior Checklist Items with Reliability Coefficients of .30 and Below and .70 and Above

9 Item-Total Pearson Product-Moment Coefficients for Checklist Items

10 Mean Z-Scores and Ranks for Independence, Dependence, and Pseudoindependence Checklist Categories

11 Ranks of Checklist Mean Z-Scores for Independence, Dependence, and Pseudoindependence

12. Weighting Matrix for Measuring Disagreement between Paired Observers
13 Distribution of Two Independent Raters' "Global" Ratings Based on Summary Checklist Data by Category and Degree
14 Frequency Distribution of Ratings by Paired Observers of Transfer Sessions
15 Summary of Agreements and Disagreements Within and Across Categories of Ratings in Transfer Sessions
16. Checklist Designations: Global Ratings vs. Z-Score Classifications
17 Global Checklist Ratings vs. Behavioral Transfer Session Ratings
18 Z-Score Checklist Designations vs. Ratings in the Transfer Situation
19 Frequency Distributions of Global Checklist Ratings, Z-Score Checklist Designations, and Behavioral Sessions Ratings (Experimental Subjects Only) by length of Exposure to Institute Program ("Filler Status")
20 Checklist and Transfer Session Ratings of Experimental Subjects and Length of Exposure to the Institute's Program ("Filler Status")
21 Chi-Square Comparisons of Ratings of Behavioral Sessions: Experimental vs. Control Subjects for Both Schools Combined and for Schools Held Separately

List of Figures

Figure

1 Geometric Design (A) for Task (2)
2 Solution for Design (A)--Task (2)
3 Geometric Design (B) for Task (2)
4 Solution for Design (B)--Task (2)
5 Partial Reproduction of Card in Set (A), Task (3)
6 Partial Reproduction of Card in Set (B), Task (3)
7 Partial Reproduction of Card in Set (C), Task (3)
8 Partial Reproduction of Card in Set (D), Task (3)
9 Cup Stimulus, Set (E), Task (3)
10 Horse Stimulus, Set (F), Task (3)
11 A Page from Booklet, Task (4)
12 Doll-Like Figure used for Demonstration Purposes, Task (5)

Chapter 1

Background and Hypotheses of the Current Investigation

General Background and Need for this Investigation

In 1962, in an effort to reverse the corrosive process of academic retardation that plagues the disadvantaged child, the Institute for Developmental Studies began to translate its prior and current research findings into action in the form of a demonstration program in Harlem public schools, carried out in cooperation with the New York City Board of Education with the support of the Ford Foundation, Office of Economic Opportunity, and other public and private funding agencies. (See Powledge /1967/ for a description of the Institute's program.)

Our enrichment program stemmed from the assumption that heavy stress on cognitive areas--specifically on verbal and language skills--would be the most effective means for intervening into the ways of thinking, learning, and behaving (with the consequent limitations in achievement) displayed by the urban, ghetto children with whom we were working. For many years, we have employed a variety of instruments for assessing the effectiveness of our work in Harlem. These have included such instruments as: The Stanford-Binet Intelligence Scale, the Peabody Picture Vocabulary Test, the Columbia Mental Maturity Scale, Reading Prognosis Test, Lorge-Thorndike Intelligence Test, Gates Reading Test, Illinois Test of Psycholinguistic Ability, the Wechsler Intelligence Scale for Children, and the Institute's specially developed Early Childhood Inventories.

As an ongoing evaluative process, the Institute continually explores various behaviors as measured by the foregoing instruments, but also, in a more general way, attempts to describe certain qualitative aspects of our program based on such sources of information as supervisors' and teachers' reports, "feedback" from parents and community observers, and principals' comments and observations. In general, findings have been most heartening, since many positive results have emerged.

Of significance and relevance to the purposes of this proposal, we should note that through the years, supervisors, teachers, visitors, and community personnel have been impressed with certain "noncurricular" aspects of our interventive impact on the children in Institute classes. Expressions such as "greater independence," "more maturity," "quieter and less motoric" have been applied to them, frequently in the context of comparisons to children in other ghetto classrooms. And indeed, it would seem that many of the teaching techniques and classroom procedures employed by the Institute are not only directed toward the teaching of content but also toward introducing changes in behavior, such as the developing of independence in the learner. Thus, such classroom techniques as the introduction of individualized activities, quiet work-time, small group involvements, decision-making activities, individual use of tape-recorder, etc. have not only been employed to teach, for example, reading, listening, and language skills, but also to aid in the development of certain behavioral and interpersonal skills. Indeed, the classroom is physically engineered to encourage the children to move about, obtain and then replace materials, and to operate equipment. We have been trying to give them opportunitites to see that they can have substantial control over their environments.

We were soon confronted with the questions of whether the Institute's enrichment classes have indeed encouraged the kinds of behaviors various observers have described. Even more important, does the behavior displayed in the demonstration classrooms "carry over"--transfer--to situations outside of the classroom?

Some Theoretical and Methodological Issues. It is important to note that several issues are relevant to our research. The first concerns the need for behavioral explorations of this sample for purposes of assessment.

(1) The need for behavioral criteria. It bears repeating, but not necessarily elaboration (since so much as already been written about this issue--see Eells, et al. /1951/; Guidelines for Testing Minority Group Children, Deutsch, et al. /1964/; Minuchin, et al. /1967/; and Zubin, Eron, & Schumer /1965/), that standardized tests, especially of the paper-and pencil variety (but also even individual tests, say, of general ability), are not culture-free. They penalize large numbers of children from disadvantaged social environments because they require skills in test-taking, listening, following instructions, and sitting still, not to mention general verbal and language ability which are simply not "built-into" the repertoire of children who otherwise may be intelligent, alert, and well able to learn. Further, from a more theoretical view, and in connection with what is emerging as methodologically desirable, we are learning that behavior is the most valid and reliable criterion for diverse measurements in personality, psychotherapeutic change, diagnosis, prediction of success in various job situations, etc., even for nondisadvantaged populations.

We note, then, that our orientation rested on the general assumption that employing aspects of behavior to predict behavior and to shed light on behavior appeared to be the most methodologically sound approach.

(2) The need for employing a "transfer model" and problems in interpreting results. Any individual working in the area of test development, with its criterion problems as well as ultimate predictive validity problems, knows (whether or not he gets around to this) that eventually his generalizations or recommendations must suffer the test of cross-validation or the test of applicability to new samples or new situations.

In the same vein, we felt that it was not enough to demonstrate that our classes produced certain types of behavior; we wished to demonstrate that this change carried over to other situations. A child "programed" by certain types of classroom environments may behave in accordance with that programing provided no change in milieu takes place. Crucial for study, not just in the proposed research, but also with regard to teaching techniques in general, was the extremely important question of transfer and generalization to other situations. Indeed, an implicit assumption of interventive education as well as conventional education is that the child generalizes what he learns (or what we think he has learned) to other classrooms, to the community, and to his daily life.

Thus, the long history in academic and laboratory circles that the transfer of training paradigm has enjoyed seemed to us more than a wasted effort in experimental minutiae. For this history offered a model for experimental exploration of a vital issue. Are we indeed making an impact which has meaning outside the teacher-classroom situation?

It should be stressed, however, that this issue is not that clear nor that easily resolved; definitive answers may not be readily available to the researcher. For intrinsic to the problem of transfer is the question of the role of situational variables. Perhaps, after all, independent behavior, at least in this age group, is partially rooted in situational and stimulus-milieu variables, and that only a careful examination of the surrounding situation (above and beyond individual personality variables) will yield accurate predictors of behavior. One implication of this issue is that failure to find transfer effects may not at all mean that we have failed to train for independence. The behavioral sessions and their tasks are, after all, rather dissimilar from the classroom situation, and the behavior elicited in them. Independence may not be, in other words, a personality trait (as noted, at least at this age level) that is "carried" within the individual from situation to situation.

Positive findings in the current study, then, will suggest transfer effects. Lack of positive findings, however, must perforce be less definitively interpreted.

Objectives and Hypotheses

As noted in the original proposal for this investigation, one of the major purposes of the current research lies in its implications for the further understanding of the independent behavior of children in and out of the classroom; an equally important purpose is relevant to assessing the effectiveness of certain Institute classroom procedures. Specifically with regard to the latter point, the basic problem under consideration is whether noncurricular aspects of behavior transfer to new situations.

It was hoped that the present investigation would also shed light on whether independence is correlated with length of exposure to the IDS enrichment classes, as well as whether there are individual variations in the extent to which independent behavior transfers to unfamiliar situations. We also hoped to ascertain whether non-Institute control Ss from the same schools could be differentiated from the Institute experimental Ss in regard to the extent to which the behaviors related to the dimensions in which we are interested would emerge in the transfer situation.

Our objectives, then, related to assessing the Institute's effectiveness in encouraging independent, autonomous behavior in its enrichment classrooms, on the general (commonly accepted) assumption that not only is such behavior developmentally and socially "normal" on the age levels with which we were working, but also indicative of the kinds of cognitive and emotional changes that underly productive, achieving future behavior--in and out of the classroom. To evaluate our success along these lines, we used behavioral rather than standardized test criteria, both in the classroom and in a "transfer" situation, to explore whether the behavior that appeared in the classroom did, indeed, emerge outside of that "programed" situation. Moreover, we wished to explore whether, in the transfer situation, relatively greater frequencies of independent behavior would be found among our demonstration pupils than among an equivalent group of disadvantaged controls.

The study, in effect, was geared to evaluate our curriculum effectiveness and goals in terms other than those related to achievement, content skills, and the like. Of significance in the design of this study is the meaningful issue of transfer of learned behavior to situations outside of the classroom with the expectation that some children in our experimental sample would not display independent behavior--a reflection of the fact that an interventive program cannot, and does not have the same impact on all members of its target population, and that there are subgroups among the disadvantaged population itself.

The general model we employed to achieve the foregoing goals was to observe (in the classroom) and then rate third-grade children from the demonstration program on systematically developed, reliable, behavioral checklists.[1]

---

[1] Equivalent control classes are in classroom situations that are conducted conventionally, with the typical classroom arrangements of seated-pupils-facing-a-teacher-at-a-desk layout and with movement, free choice of materials, etc., not encouraged. Such Ss could not, perforce, be rated and compared to the demonstration Ss.

The latter instrument-development procedure proved to be a major aspect of the year's work. Two trained, independent experimenters observed the experimental Ss in the classrooms over a period of time, employing the behavioral checklist. Two new raters then anonymously rated the material based on behavioral observations of the children along a four-point discontinuous scale: independent (productive); pseudoindependent; nonindependent; and can't rate, don't know. (The "pseudoindependent" classification derives from the basic fact that behavior that subsumes motility, detachment, noninvolvement, inattentiveness, and the like--might be rated as independent, but that such behaviors could hardly reflect the construct in which we were interested.) Following a behavioral, non-classroom observed transfer situation which allowed independent behavior to emerge (and the development of transfer tasks also proved to be a major research activity), observers who did not observe in the classroom rated random small groups of demonstration and equivalent control subjects on the same four-point discontinuous scale. No rater had any prior knowledge about any S.

Our hypotheses and expectations based on the above procedures (the latter are described in complete detail in the next chapter) follow:

(1) After operationally defining so-called independent, mature (age-related) behavior, we hoped to be able to distinguish the behavioral variables associated with this attribute from "independent" behavior which is not productive or useful; and these behaviors, in turn, from nonindependent behavior. We expected a range of such behaviors in the enrichment classrooms with a relatively larger number of "independent" ratings than the other types of ratings; and, we also expected that F3 "filler" Ss (replacements due to attrition who entered our program in the third grade), since they had virtually no years of exposure to the enrichment program, would be rated relatively less frequently as "independent" than the other experimental Ss.

(2) Having thus identified behavioral subgroups along this dimension in the demonstration classes, we expected that in a non-classroom, behavioral transfer situation, those rated as "independent" would tend to receive the same rating by another team of observers who had no prior knowledge about the preceding ratings; the same expectations were held for the Ss rated as "pseudoindependent" and "nonindependent." In other words, we expected our prior classification along this trichotomy to be related to classifications of the same Ss as observed and rated in a new situation. "Filler" status would of course be considered in this analysis.

(3) We expected that control Ss would receive fewer ratings of "independent" in the transfer situation, and relatively more ratings of "dependent" than the demonstration Ss, again, with "filler" status of the experimental Ss considered.

(4) Needless to say, underlying the past year's research activities was the basic expectation that we would develop a reliable, usable classroom checklist, which we could later employ in more theoretical explorations of independence along construct validity lines.

(5) A subsidiary expectation concerned the notion that on the basis of checklist designations, trained "global" raters (without seeing the Ss) would rate some of them as "pseudoindependent," and that the latter is a meaningful category in this regard.

Finally, we should add that we were quite interested in exploring the role of situational variables in determining behavior, on an assumption that independence is not, after all, "carried within" the individual from situation to situation regardless of the nature of that situation.

Chapter 2

Sample and Methodology

The general model for the currently described research was briefly presented in the preceeding chapter. The design included: (a) the development of a reliable behavioral classroom checklist for independent, pseudo-independent, and dependent behavior; (b) the formal phase of classroom observations (over time) using the checklist (involving two observers); (c) the "global" rating of these checklists by two different independent raters; (d) the development of quantitative, continuous scores based on checklist observations; (e) the development of behavioral tasks for the "transfer" sessions which would yield an opportunity for reliable ratings to be made; and (f) the "running" of experimental and control Ss in these transfer sessions with independent observations and ratings made by pairs of observer-raters. Note, the first (classroom) phase of the research utilized children (the experimental Ss) from the Institute's classes. Non-Institute classes were not observed for independent behavior, since, in terms of their physical layout, they did not offer an opportunity for relevant ratable behaviors to emerge. The behavioral sessions, on the other hand, were based on small groups of randomly assigned experimental and control children. The Institute children had, it should be noted, varying degrees of exposure to our enrichment classes.

The current chapter describes: (a) the subject population with which we worked in this investigation; (b) development and description of the classroom behavior checklist; (c) development and description of the behavioral tasks for the transfer sessions; and (d) description of the global rating employed in the various phases of our study. Chapter 3 presents information concerning the development of other scores, the checklist revision, various reliability explorations we conducted, and data analyses and findings for various parts of the study.

Sample

Experimental Ss--Phase I (Classroom Observations)

The larger population from which subjects for this study were drawn consisted of third graders (1968-1969) in the Institute's demonstration classes in Public Schools 68, 79, 90, and 175 in Harlem.

Because the currently described study was of only one year's duration, we could not work with third graders from all four of the foregoing schools. Had we done so, about 150 subjects would have been involved when equivalent controls were added. Since this was not a feasible sample size, two schools were selected on the basis of their yielding the largest numbers of third-grade children who have had at least two years of exposure to the Institute's enrichment program.

In the classroom observation phase of the research, 19 children were observed at Public School 90 and 21 children at Public School 175. Of the 40 children included in this sample, 28 children had at least two years of exposure to the Institute's program (12 at Public School 90, and 16 at Public School 175). The distribution of this sample by school, age, sex, and length of time in the program is shown in Table 1.[2]

Table 1 shows that average age for each school is virtually the same, and that only minor age differences for the various types of subjects within each of the schools are found. Most Ss, as of September, 1968, were eight years and a few months old; a few Ss were just under eight years of age; one S was seven years and four months, and another, seven years and five months old. Sex distribution is uneven in this population, with almost two-thirds of the total sample consisting of boys.

Since years of exposure to the IDS program is a variable in which we are interested, and since the development of independence is age-related, it is interesting to see whether or not wide differences in mean ages are found between subjects classified as to "filler" status. Table 2, which presents mean age for various classifications of Ss (with both schools combined), indicates that as of September, 1968, average age for each classification is approximately 8 years and a few months, with wide differences among groups not found.

---

[2] The first progress report (December, 1968) described a sample of 41 children (see Table 1 in Progress Report #1). One S at Public School 90, however, moved out of that school district before the classroom observations could be completed.

Experimental and Control Ss--Phase II (Transfer Sessions)

The Ss described above were observed along with control Ss in the behavioral transfer sessions which will be described shortly. The actual experimental sample was reduced by one because a female S from Public School 90 was not available when the behavioral transfer sessions were run. She was an F3 S, thus reducing that group from 9 to 8 Ss and the total experimental sample in the transfer sessions from 40 to 39. Table 3 presents the distribution of this sample by school, age, sex, and length of time in the program, and Table 4 presents mean age for these classifications with both schools combined. When Table 4 is compared to Table 2, it can be seen that only minor changes were introduced by the elimination of one S.

Control Ss were drawn from the third-grade classes of Public Schools 90 and 175. No selective factor was introduced: controls were taken from all the non-Institute third-grade classes in these schools. Table 5 presents mean age and sex breakdown for these subjects. Note, although 39 control Ss were actually "run" in the behavioral sessions, only 37 were regarded as control Ss due to sampling errors.

The control Ss come from the same SES, school, and community backgrounds as the experimental Ss. Mean ages are almost identical with those of the experimental Ss, it should be noted.

The Classroom Behavior Checklist

This section describes the development of the Classroom Behavior Checklist for Dependent, Independent, and Pseudoindependent behavior, as well as the checklist itself as we actually used it in initial phases of the current investigation.

Working Definitions of Behavioral Categories

During the early periods of pilot work, when we were considering items for checklist inclusion, we clarified our thinking concerning what we meant by the various behavioral categories. It should be stressed that this clarification took place within the context of constant feedback from our research team as to their observational experiences with groups of children using some of the earlier items. It was during this early period, for example, that we confirmed our expectation that our prior classification, pseudoindependent, had some behavioral meaning. (We saw, however, that this term could be ambiguously interpreted, since it seemingly did not imply a behavioral referent other than that which could be regarded as a residual or derivative from the independent behavior category.) Raters, teachers, supervisors, and others on our professional staff all seemed to agree that certain behaviors as observed in the classroom fell neither into an independent nor a dependent category, but yet were still classifiable--that is, such behaviors were easily identifiable and were not of a "don't know" nature. Indeed, this third category (but not middle category, since a continuum is not implied) subsumed behaviors which our observers suggested were rather salient and recognizable, for their characteristics were overt and often possessed an attention-getting quality.

From sources in the literature, our own experiences in the schools, and feedback from our research team, then, the following working definitions of the behavioral categories emerged:

(1) Independent behavior is oriented toward productive (constructive) goals, is age-related and is self-initiated. The independent child manifests behavior which is both self-reliant and self-assertive. He is considered independent of others to the extent that he attempts to satisfy his needs in a manner which is not entirely a function of the way others behave toward him. A child is considered independent when he conducts his problem-solving activities without seeking unnecessary help from others, and is able to carry on his activities without continuous need for reassurance, affection, or approval. Independence also includes behavior which is self-assertive, such as striving to master tasks. Independent children characteristically show relatively good motor control, that is, their motility is appropriate to the situation; they are neither overly constricted nor excessively motoric.

(2) Pseudoindependent behavior can be described as nonproductive, nondependent behavior, characterized by the seeking of attention through asocial or antisocial activities, the maintenance of physical and psychological distance from peers or teachers (leaders), high motility, and inconsistency in the pursuit of goals. Children so designated are likely to interfere with the activities of other children, refuse to participate in group activities, and be destructive of work materials and equipment. Such children often do not accept help or follow instructions, even though their own efforts are clearly not successful. They also tend to show random sequences of behavior with unrelated activities begun and discarded before completion.

(3) Dependent behavior is determined by the presence of specific characteristics, not merely by the absence of characteristics specific to independence. Responses that are characterized as dependent, are often those that seem directed toward obtaining social reinforcement. Such behavior includes many different kinds of activity, such as: persistently seeking help (whether necessary or unnecessary); affection-seeking; and bids for

attention, reassurance, approval, succorance, and guidance. Dependent behavior is also manifested in passivity and lack of spontaneity.

Development of the Classroom Observation Checklist

Early in September, 1968, various classroom behavior scales and reports in the literature were examined for the construction of a large pool of possible items. Some of these sources were: Beller (1955); Cervenka (1967); Davidson and Greenberg (1967); Finney (1961); Gewirtz (1956); and Hartup (1963). As noted, this procedure was also combined with group discussions and feedback concerning refinement and clarification of what was meant by each of the categories (independent, dependent, and pseudoindependent). Later that month, each item was discussed from the point of view of a minimal standard--face validity. At this point, it became clear that items reflecting motor activities (such as hyperactivity in the classroom) and distractibility should be included in the pool of items; it also became clear that there were many seemingly distinct items which were, in fact, reflecting identical behavioral elements. Gradually, more items reflecting the child's interaction with other children were developed. We also began to consider, at this time, the mechanics of the actual observations.

Pilot-testing was begun in October, 1968, using second-grade children in Public Schools 175, 90, and 79 in Harlem. Third-grade children at the Friends Seminary were also observed. The original checklist was revised three times on the basis of pilot observations. During this preliminary period, the observers worked with and evaluated the feasibility of using various observation units (that is, size of units or groups of children who were observed), durations of observation periods, and other classroom observation procedures.

In the light of certain difficulties encountered in observing the children in small groups, or in too short segments of time (the behaviors in which we were interested required some continuity of observation-over-time), as well as because of other mechanical considerations (for example, the unwieldiness of checking the behavior of several children at a time on large checklists), the procedure finally adopted in November was: two observers independently observed each child in a single classroom for five-minute periods, in sessions lasting up to two hours. The particular hours selected were to be scheduled well in advance and took into consideration the teacher's daily schedule of classroom activities. (We learned, for example, that observing children cooking, or having juice, or dancing in a group did not yield the kind of variation in behavior which we were seeking.) A brief "narrative" was written by the observer as he was watching the particular child. At the end of each day of observation, each observer independently reviewed his own narratives, and the relevant behavior was checked on the checklist (for each child) along a four-point scale--frequently, moderately often, rarely, and never. A fifth category, no opportunity to observe, was also included.

The checklist at this point had been revised to include nine items in each of the three behavioral categories (see next section). By about the middle of November, the checklist was tested once more (Public School 175) by having each observer observe and rate each child for a five-minute period. The two observers showed high reliability in their observations.

The formal data collection period began in November, 1968, and ran through most of January, 1969, in Public Schools 175 and 90. Subjects were independently observed by two observers in their normal classroom settings. The total number of sessions and observations per child ranged from fourteen to twenty-one, with a median of twenty. No difficulties in either using the checklist or in the actual mechanics of classroom observations arose during the testing period.

The schedule for both observers was based on random selection of days of the week and random time-cluster sampling at each school (modified only to balance the number of observations at each school). In addition, separate randomized lists of subjects were prepared for each observer for each day of observation.

Departures from strict randomization of the ordering of subjects were necessitated, however, by the absence of some of the children for various reasons (e.g., some were not in school on a particular day, some were sent on errands, some were taken out to rehearse their parts in a forthcoming play, and sometimes the group activity so structured the situation that valid observations could not be made, as in story-telling time).

Whenever possible, a child who was not available at the scheduled time was observed later on a randomly determined basis. If a child was absent more than four times, the observation period was extended so that the ratings were finally based on a minimum of 16 (with a maximum of 20) observations on each child for each observer.

At the end of each day, all identifying data were removed from the narratives after they were typed. At the end of the several-week observation period, each observer independently summarized his daily checklist on one master list for each child, with frequency counts summarized for each category for the four-point scale: frequently, moderately often, rarely, and never (as well as no opportunity to observe).

Note, narratives allowed the opportunity for singular or unusual events to be recorded--events that were not anticipated, or even though they occurred but once, were thought to bear some significance to what we were after. Narratives also could provide, we thought, an opportunity to qualify a behavior tally on the checklist--for example, a child who was considered unusually hyperactive was later taken home by his mother because the nurse reported that he had a high temperature. This procedure was introduced in the light of the possibility that some unique aspect of behavior or some special circumstance could arise which was not or could not be anticipated by the behavior checklist, despite our careful pilot-testing in many classroom situations. Such narratives, we expected, would allow the final raters to take into consideration a qualifying, explanatory, or modifying circumstance which might alter the picture obtained only from frequency counts based on the checklist. In actuality, our classroom observers have noted that only in rare instances did such narratives "clarify" the checklist information, and they were of little or no help when the final global ratings were done. Further analysis of the narratives was therefore not undertaken.

The Classroom Observation Checklist Items

Appendix A presents the behavior checklist as employed by two independent observers of all the children in the same classroom situation (at different times during at least a two-hour period at each school) on different days over a period of almost two months. The sequence of items is a random one. Below are the checklist items (with examples of actual behavior) regrouped into categories according to what we think they are reflecting, as judged by construct considerations, as well as empirical considerations during pilot phases of our exploration.

Independence Items--9

(4) Volunteers contributions to class discussions and projects--e.g., offers information without being called upon.

(5) Helps, sympathizes, shows consideration of others, is thoughtful, offers praise, guidance, assurance--e.g., helps other with work when appropriate, lends or shares materials freely.

(6) Sits quietly and attentively when task requires--e.g., remains undistracted by irrelevant activities.

(9) Utilizes available free time constructively--e.g., initiates productive activity when there is no ongoing class activity.

(11) Withstands interference while engaged in individual activities, as in doing difficult assignment, puzzle, painting, etc.--e.g., does not abandon task when interrupted.

(14) Takes appropriate initiative in problem-solving situations--e.g., asks questions for information which seem to be prompted by general curiosity rather than bids for attention; seeks needed help; asks questions in order to obtain further information or clarification for task at hand.

(20) Tries to figure out things for himself before calling on teacher (leader) or other children for help--e.g., tries to overcome obstacles in the environment.

(21) Enters readily into new situations; is venturesome, inquisitive, etc., when appropriate to ongoing activities--e.g., is eager to begin new and different activities.

(25) Attempts to resolve difficulties that arise between himself and other children without appealing to teacher (leader)--e.g., does not initially seek adult intervention to solve difficulties between himself and peers.

Dependence Items--9

(1) Withdraws in the face of difficulty--e.g., abandons goal when presented with an obstacle.

(12) Imitates others rather than initiating own activities--e.g., changes own behavior to conform to the influence of other children.

(13) Gets upset when called upon in class; appears fearful, tense, timid--e.g., requires repeated urging before making a verbal response.

(15) Seeks to be near others--e.g., follows teacher, assistant, leader, classmates about the room.

(17) Requires close assistance or direction in order to work at a task--e.g., easily distracted; must have teacher or assistant's help to remain attentive to task.

(19) Lets other children impose on him or boss him around--e.g., does not stand up for his rights; yields his place in line, relinquishes turn at play, etc.; allows others to exploit him.

(23) Seeks constant recognition and/or approval--e.g., tries to get teacher's or assistant's attention as frequently as can.

(26) Approaches new tasks timidly and without confidence, shrinks from trying new things--e.g., requires strong encouragement before participating in activities.

(27) Tends to choose habitual and familiar activities.

Pseudoindependence Items--9

(2) Is inattentive, is easily distracted by things going on around him--e.g., engages in random class-unrelated activity (wanders about, daydreams when should be working on own); talks to visitors about irrelevant matters when he/she should be engaged in a task; daydreams; attention wanders from task at hand; is not prepared to answer when called upon in class.

(3) Uses materials and equipment in a rough or destructive manner--e.g., destroys books, crayons, games etc.

(7) Brags, shows off, displays exaggerated opinion of own abilities.

(8) Seeks attention, i.e., in a manner that seems designed for active notice rather than stemming from a need for help--e.g., is inappropriately aggressive in seeking the attention of adults.

(10) Interferes with the work of others--e.g., imposes on other children and bosses them around; inappropriately insists on being admitted to a small peer group; destroys or takes away the work or materials of other children.

(16) Does not follow directions--e.g., works or plays by self despite teacher's request that all children engage in group activity.

(18) Seeks to do things differently from others, even when own method is not effective--e.g., persists at tasks which are clearly impossible.

(22) Does not pursue task to completion--e.g., leaves task at hand

for another or to wander about the room.

(24) Speaks out of turn; interrupts others; takes others' turn, etc. --e.g., refuses to wait his turn.

The Behavioral Transfer Sessions

We had originally planned to eliminate from the second phase of the study any children for whom unambiguous ratings of Independent, Dependent, or Pseudoindependent could not be made. Since the two independent raters had not completed their ratings of the Ss (on the basis of data collected in the classroom observations) and the Z-score checklist designations (see next chapter) also were not completed by the time scheduled for running the transfer sessions, the entire sample was included in the second (transfer) phase of the study. In the final analysis of data, however, all children for whom it was not possible to make an unambiguous rating were excluded.

Transfer session leaders and observers were trained during pilot-testing of the tasks on second-grade children at Public Schools 79 and 175. The leaders were two experienced psychological testers (one male and one female). Each leader conducted approximately one-half of the experimental sessions at each of the two schools from which the control and experimental samples were drawn (Public Schools 90 and 175).

Four experienced observers were also trained, two to serve regularly, and two to substitute on occasions when it was necessary to schedule sessions simultaneously at both schools. In the latter case, one of the regular observers was always paired with one of the substitute observers.

The two session leaders and the four observers were staff members who were not involved in the collection of data in the first phase of the study; each, of course, did not know the composition of the groups, that is, the experimental or control status of the subjects.

The leaders and the observers were rotated between the schools, and the children were randomly selected for each session at each school.

As noted, there were 39 experimental Ss. Since 39 Ss were also run as controls, the total number of children in the experimental sessions was 78. There were, in all, 15 sessions: seven sessions at Public School 90 (involving 36 children); and 8 sessions at Public School 175 (involving 42 children). Table 6 presents the distribution of experimental and control Ss by sex and school for the transfer sessions.[3]

Parental permission was obtained for working with the control children in this study. The principals of the schools were extremely helpful in guiding us in this procedure to insure the greatest degree of co-operation from parents and teachers of the non-Institute children.

Subjects were run in groups of four or six, the composition of each group being consistently one-half experimental and one-half control children. Other than this criterion, the children were completely randomized, so that the mixture of independent, pseudoindependent, and dependent Institute children varied from group to group. This procedure was adopted to eliminate the possible experimenter bias that might emerge if the observers knew that the composition of the group was always partially structured by the distribution of the experimental Ss.

The original plan had been to match experimental and control Ss by sex. This, however, proved not to be feasible because of difficulties encountered in obtaining parental permission for use of control Ss. Ultimately, to avoid causing excessive hardship to school personnel, it was necessary to drop the requirement for matching control to experimental Ss by sex.

Development of the Transfer Tasks

We devoted considerable staff discussion during the early part of the year to the exploration of various possibilities for tasks, games, puzzles, and tests to be used in the transfer situation which would allow a variation of behavior to emerge. As a matter of fact, members of our research team made several trips to various commercial dealers of toys and playthings in New York City to obtain ideas for techniques, puzzles, games, etc, which might be appropriate for our needs. Specifically, Creative Playthings and Childcraft were visited as well as several department stores.

---

[3] See Note, Table 5.

The selection of such tasks was by no means a simple procedure. The tasks had to be difficult enough, with children of unequal intellectual capacities, to allow choices for each child in terms of seeking help, asking appropriate questions, working diligently on his own, seeking help even though the solution is reachable, etc. Tasks that were too easy would not allow dependent behavior to emerge. Tasks that were too difficult would cause all children to seek help or to give up. Further, whether the Ss would be observed working together in small groups or alone as they solved the tasks had not yet been resolved.

By the end of October, we had considered, but had neither accepted nor rejected, various tasks and situations including, among many others, the following:

(1) An Asch-type task (Asch, 1956) which would allow conformity-dependency behavior to emerge. In this situation, Ss would have the opportunity of producing written or verbal responses, so that objective measures of conformity-dependency would be available.

(2) Various standardized or semistandardized tests and clinical instruments such as the Block Design test, the Goldstein-Sheerer tests, the Vigotsky test, the Stencil Design test of the Arthur Point Scale, etc.

(3) Various experimental techniques, such as the level of aspiration procedure so popular in the 'forties, in which a child could be induced to conform to, or could withstand, influence to alter his level of aspiration

in accordance with the examiner's instructions (or quasi-information about the subject's performance or the performance of peer groups).

(4) A large jigsaw puzzle of a map of a moderately unfamiliar area, possibly Canada. With this task, the Ss could be instructed to put the pieces together as best they could, on their own. They would be informed, however, that there was a large map of Canada behind a screen which they could consult, if necessary. Persistence and the lack of consultation would be the criteria for best (independent) performance.

The extensive pilot-testing of the tasks to be used in the transfer sessions was based on many of the same second-grade children who were observed during the preliminary work with the behavior checklist, thereby giving us the advantage of knowing their characteristic classroom behavior.

The tasks for the transfer sessions that were finally developed to elicit behavior relevant to the three categories in which we were interested are described in the next section. One of the major criteria for their selection was the range or variety of behaviors they seemed to elicit, including interaction with peers as well as authority figures. In the pilot phases, we also evaluated various tasks from the point of view of ease of task-administration. Finally, during this preliminary period, we assessed the potential of various tasks to elicit rater-agreement as to the Ss' behavior.

We devised some unstructured as well as some structured situations to permit different types of behavior to emerge. The structured situations were generally based on the Asch model (1956), modified to avoid the need for a confederate or "stooge" role. This was necessary since it was not feasible to use the children as confederates. The use of an adult confederate would have confounded the situation; Crandall (1967), for example, reported that earlier research had indicated that compliance to adult pressures and values was found among high-achieving children of this age group.

The behavioral tasks, which we devised for peer groups of four or six Ss, are described below, as noted. It will be seen that tasks (1), (3), and (4) are relatively structured, while tasks (2) and (5) are relatively unstructured. The latter tasks provide opportunities for moderately wide variations in behavior, while the former do so to a lesser extent.

In the actual administration of the experimental sessions, tasks were presented in rotation so that the order in which they appeared was balanced across all groups. Each session lasted for approximately three-quarters of an hour.

The behavioral sessions were conducted in vacant classrooms in Public Schools 90 and 175. Desks and chairs were arranged in the front of the room so that the subjects could be seated with empty desks between each of them. The raters sat at either end of a large desk placed at the side of the room. Here they could easily observe the children without being in their direct visual field. The materials used for the various tasks were also kept at this desk. Another small desk was situated in front of the room, directly under the blackboard. This desk and the chalk tray were used for displaying materials to the children.

Prior arrangements had been made with the teachers for taking the children from their classrooms. While the leader went to pick up the children, the raters prepared the room and task materials. This procedure insured that the raters remained unaware of which children were experimental Ss and which were control Ss.

During the sessions, the group leader stood at the front of the room, attempting to create as informal an atmosphere as was possible while still maintaining the subjects' attention. As the sessions progressed, the two observer-raters recorded frequencies of various behaviors as well as comments on worksheets which had been developed and polished during the pilot phases. These observations, special comments, responses, and tallies were used by the observers in their ratings of each *S*. Immediately after each session, then, the observers independently rated each *S* on the following scale:

Rater_______________________ IDS Code No._______________________

Check only one:

I--Independent	Highly______	Moderately______
P--Pseudoindependent	Highly______	Moderately______
D--Dependent	Highly______	Moderately______
X--Can't rate		

Degree of confidence in rating--check one:

3____high	2____medium	1____low

It should be noted that all tasks were employed in qualitative fashion by the raters who, in effect, were rating the *S*s globally.

Administration and Description of the Tasks

After all the subjects were in the room, *E* said:

> GOOD MORNING (AFTERNOON). I'M MR. (MISS)_____________________.
> NOW I'D LIKE EACH OF YOU TO INTRODUCE YOURSELF.

E called on each *S* to give his name and recorded each name for identification purposes. *E* then said:

> YOU'RE GOING TO BE WORKING ON DIFFERENT KINDS OF GAMES AND THINGS. WHILE I"M GETTING THE THINGS READY, YOU MAY COME AND GET A BOOK TO READ, IF YOU WISH. YOU MAY TAKE IT BACK TO YOUR SEAT.

E pointed to a large stack of books. This introduction helped to ease the *S*s into the situation and also occupied them while *E* was preparing the materials.

(1) Peer influence--bean guessing.

Jars that were approximately 1½" in diameter, ranging in height from about 2" to 7" and filled with beans were presented in ascending order of size. There were six such jars available, but either four or six were employed depending on the size of the group (i.e., either four or six Ss).

> ALL RIGHT. WE WILL PUT THE BOOKS ASIDE NOW, AND TRY SOME GUESSING GAMES. HERE IS A JAR FILLED WITH BEANS. I'M GOING TO CALL ON EACH OF YOU, IN TURN, TO GUESS HOW MANY BEANS THERE ARE IN THE JAR. NOW, DON'T ANSWER UNTIL YOU'RE CALLED ON. __________ (CHILD #1): HOW MANY BEANS DO YOU THINK THERE ARE IN THIS JAR?

E then called on each child in order.

> VERY GOOD. NOW, HERE IS ANOTHER JAR FILLED WITH BEANS. LET'S SEE HOW WELL YOU CAN GUESS ON THIS ONE. REMEMBER, DON'T TALK UNTIL I CALL ON YOU. THIS TIME __________ (CHILD#2) WILL START. ________, HOW MANY BEANS DO YOU THINK THERE ARE IN THIS JAR?

This procedure was repeated with the jars of different sizes until each of the children had responded in all order positions.

This behavioral situation, an Asch-type structured task, indicated the extent to which S's guesses were influenced by those preceding him; or, conversely, the extent to which his judgments remained independent of guesses made by other children in the group.

(2) Spontaneous behavior, self-reliance, peer-influence--geometric puzzles.[4]

For this task, a seven-piece, colored plastic puzzle was employed. A great variety of geometric designs can be constructed with this puzzle by arranging the pieces in various ways. Although the puzzles appear simple, they are in fact quite difficult to solve. The two designs selected for our task are shown in outline form in Figures 1 and 3, while their corresponding solutions appear in Figures 2 and 4. All figures are scaled to actual size. The pieces of each puzzle distributed to the Ss were numbered with a marker to correspond to the numbers on the solution diagrams. E introduced this task by saying:

---

[4] This is a Kohner Educational Puzzle, called VOODOO, manufactured by Kohner Bros., Inc., P.O. 294, East Paterson, N.J.

> OKAY, NOW WE'RE GOING TO DO SOMETHING ELSE. I'M GOING TO GIVE EACH OF YOU A PUZZLE, AND I WANT YOU TO MAKE A DESIGN. TAKE THE PIECES OUT OF THE ENVELOPE, BUT DON'T START UNTIL I EXPLAIN EVERYTHING.

While talking, E distributed the envelopes, placed the outline of the first figure on a stand in front of the Ss, and the solution face down on the desk. Pointing to the displayed outline, E said:

> NOW, TO MAKE THIS FIGURE CORRECTLY, YOU MUST USE ALL SEVEN PIECES. TRY TO WORK THE PUZZLE AS FAST AS YOU CAN.

E briefly flashed the solution:

> I'M GOING TO PUT THE PICTURE THAT SHOWS HOW TO WORK THE PUZZLE HERE. I WANT YOU TO TRY YOUR BEST TO DO THE PUZZLE YOURSELF--BUT, IF YOU REALLY THINK YOU CAN'T DO IT, YOU MAY COME UP AND LOOK AT THE PICTURE. OKAY, START NOW; AND WORK AS FAST AS YOU CAN. TRY TO DO THE VERY BEST YOU CAN.

E withdrew to the back of the room for several minutes, and then came forward to ask:

> WOULD ANYONE LIKE SOME HELP WITH THE PUZZLE?

To maintain the children's interest and give them a sense of closure, E made sure that everyone had succeeded in solving puzzle (A) before the second design (B) was introduced. E then said:

> NOW SCRAMBLE ALL THE PIECES OF THE PUZZLE AND WE WILL TRY TO MAKE ANOTHER ONE.

The same procedure was followed for the second design (B).

This task provided such information as: persistence before looking at the solution; whether the child copied from his neighbor in attempting to solve the puzzle; whether the child really worked at the solution or simply did not try; whether the child imitated another child by going up to look at the solution; and whether the child resisted or readily accepted the offer for adult assistance.

(3) Peer influence--circle, square, star, dot guessing.

Six sets of six cards each were prepared:

Set (A) was constructed of black posterboards 11" x 14". Each card contained approximately 150 stick-on circles of various bright colors. For easy identification, the cards were labeled with the letters F, H, J, L, N, and R. Figure 5 is a partial reproduction of a card in this series.

Set (B) consisted of 11" x 14" blue posterboards each containing stick-on circles of a single color: either green, orange, pink, red, tan, or yellow. Figure 6 is a partial reproduction of the green card.

Set (C) also consisted of 11" x 14" posterboards, three with black and three with grey backgrounds. Each card was pasted with squares of a single color, either blue, green, orange, purple, red, or yellow. Figure 7 shows part of a card from this series.

Set (D) was constructed of gummed stars of various colors on 9" x 11" posterboards. The backgrounds of these cards were red, green, or blue and there were two cards of each. These cards were labeled G, L, O, S, X, and Y. See Figure 8 for an example of this set.

Set (E) contained 8-inch square drawings which were used in another (unrelated) research investigation with children.[5] These consisted of a heavily dotted background containing a more finely dotted figure of an object: clock, cup, gun, hand, key, and safety pin. The cup stimulus is shown in Figure 9.

Set (F), also developed previously, was similar to Set (E) except that the background consisted of a grid instead of dots. The common objects pictured in this series were: boat, horse, pail, pan, rabbit, and scissors (see Figure 10, the horse stimulus).

---

[5] C. Deutsch and F. Schumer. Brain-damaged children: a modality-oriented exploration of performance. New York: Brunner/Mazel, Inc., in press.

Arranging Set (A) for display, E said:

> (A) ALL RIGHT. NOW WE'LL TRY ANOTHER GUESSING GAME. I'M GOING TO PUT THESE CARDS HERE, AND I WANT EACH OF YOU TO GUESS WHICH ONE HAS THE MOST CIRCLES. DON'T ANSWER UNTIL I CALL ON YOU. WHAT I WANT YOU TO DO IS TO GUESS WHICH CARD HAS THE MOST CIRCLES--CARD F, H, J, L, N, OR R.

E called on each child in a predetermined order. He then introduced the five (or three) additional sets of stimuli, calling on the children in a predetermined order so that each child responded in all order positions. (All six sets were used when a full group of six children was present. When the group contained four Ss, only the first four sets of cards were introduced.)

> (B) OKAY, NOW WE'LL TRY ANOTHER GUESSING GAME, WITH SOME DIFFERENT CARDS. EACH OF THESE CARDS HAS CIRCLES OF A DIFFERENT COLOR. BE SURE NOT TO ANSWER UNTIL I CALL ON YOU. WHICH ONE HAS THE MOST CIRCLES: GREEN, ORANGE, PINK, RED, TAN, OR YELLOW? THIS TIME ______________ WILL START.

All other sets were introduced with instructions similar to those for (B), except that the Ss had to guess about squares, or stars, or dots.

This is also an Asch-type task, giving indications of conformity to peer influence.

(4) Self-reliance--counting animals.

For this task, 8½" x 11" booklets with drawings of animals and other common objects were constructed. Each page of the booklet contained nine pictures. The total number of animals pictured in each of the six booklets ranged from 48 to 53. Thus, each S would arrive at a different number if he counted the animals correctly. Figure 11 is a page from one of the booklets. E said:

> NOW WE'LL DO SOMETHING ELSE. AFTER I FINISH EXPLAINING YOU CAN START, BUT DON'T START UNTIL I TELL YOU WHAT I'D LIKE YOU TO DO.

E distributed a booklet to each child, face down. E continued:

> I WANT YOU TO GO THROUGH THE BOOKLET AND COUNT ALL OF THE PICTURES OF ANIMALS THAT YOU SEE THERE. IF YOU FINISH COUNTING BEFORE EVERYONE FINISHES, JUST LOOK UP, BUT DON'T SAY ANYTHING. SIT QUIETLY AND WAIT UNTIL EVERYONE FINISHES, SO THAT YOU WON'T INTERFERE WITH THEIR COUNTING. AFTER YOU HAVE ALL COUNTED THE ANIMALS, I'LL ASK YOU TO TELL ME HOW MANY YOU COUNTED.

After all the children had finished counting, E called on them, one at a time, for their estimates. E then continued:

> NOW I WOULD LIKE TO BE SURE THAT YOU COUNTED CORRECTLY, SO IF YOU DON'T TRUST YOUR OWN COUNTING, YOU MAY ASK SOMEONE ELSE TO CHECK IT OVER FOR YOU.

This task is another variation of the Asch-type situation, intended to elicit measures of self-reliance.

(5) Authority and/or peer influence--picture construction.

The materials for this task consisted of sheets of colored geometric "stick-ons" and 5" x 8" index cards. A simple doll-like figure using a few of the stick-on shapes given to the children had been constructed for demonstration purposes (see Figure 12).

Displaying the preconstructed design, E said:

> NOW I'M GOING TO GIVE EACH OF YOU A CARD AND SOME OF THESE COLORED SHAPES, AND YOU CAN EACH MAKE A PICTURE. JUST TO SHOW YOU HOW THESE SHAPES CAN BE USED, HERE IS A PICTURE SOMEONE ELSE DID. YOU CAN MAKE ANY PICTURE YOU LIKE. WHEN YOU'RE FINISHED, TURN THE CARD OVER AND PUT YOUR NAME ON IT. I WANT TO KEEP THE CARDS FOR AWHILE, BUT THEN I'LL GIVE THEM TO YOUR TEACHER SO THAT SHE CAN RETURN THEM TO YOU.

This task was designed to indicate S's conformity to authority and peer influence. The Ss could either duplicate the E's picture, copy from their neighbors, or construct their own pictures. The pictures were returned to the Ss after the year's work was completed to avoid contamination of responses to this task.

The next chapter presents the findings based on the research methods described in the current chapter.

Chapter 3

Findings

The material in the present chapter is presented, roughly speaking, in two sections--reliability findings and the development of certain scores; and results of data analyses pertaining to the testing of hypotheses.[6]

I

Reliability Measures and the Development of Scores

This section consists of the following: (a) checklist--reliability of observations for paired observers for each item; (b) checklist--item-total correlations (within each category); (c) checklist--the development of quantitative, continuous scores yielding subject classification procedures (based on Z-scores); (d) reliability of "global" ratings based on checklist--two independent raters; (e) reliability of ratings in the transfer sessions--two independent raters; and (f) checklist--the relationship of Z-score designations and "global" ratings.

Checklist--Reliability of Observations for Paired Observers for Each Item

As will be recalled, and as can be seen in Appendix A, the Classroom Behavior Checklist comprised items which fell into the three categories of independence, pseudoindependence, or dependence. Observations of these behaviors for each day of observation for each child (over a period of time) were tallied for the different frequency points: frequently, moderately often, rarely, and never.

To calculate the reliability of observations per item, for each S a frequency count was made of the number of tallies per item for each of the foregoing frequency points for each of the observation sessions, with data for each observer treated separately. The can't rate point was excluded in this procedure. Totals of these counts for each child per item across all of the observation sessions were then calculated. These computations yielded the total number of "ratable" time-sampling observations made for each child for all sessions for each observer. To obtain a score for each child, weights were assigned: frequently--3; moderately often--2; rarely--1; and never--0. (Higher scores for any item reflect "more" of the particular behavior observed, i.e., greater independence, greater pseudoindependence, or greater dependence.) These weights were then multiplied by the number of observations for the scale point, yielding a total score for that scale point. These scores were then totaled across scale points. Totals were then divided by the number of ratable observations to yield a mean score for the child, thereby eliminating irrelevant variance resulting from unequal numbers of observations. These item total scores were finally transformed by adding a constant (+1) to each, to eliminate zero scores and products in determining inter-observer reliability.

---

[6] See Edwards (1954) for a description of the various statistical procedures we employed.

The total scores for each item were paired for each independent observer, and Pearson correlation coefficients were then computed for these paired scores. The coefficients thus obtained indicate the amount of agreement between observers. These correlation coefficients are reported in Table 7.

An examination of Table 7 discloses a wide range in magnitude of the obtained coefficients--with the pseudoindependence category containing the largest number of items with higher reliabilities. Because reports by the classroom observers indicated extreme difficulty in working with some of the items (some relevant behaviors simply did not occur), we decided to systematically explore the possibility that the items with lower reliability coefficients were those which were simply not ratable or for which there were no opportunities to observe relevant behavior, while the items with higher coefficients were those which related to behaviors which occurred with greater frequencies. Accordingly, we arbitrarily established coefficient cut-offs of .70 and above and .30 and below, and tallied frequencies of the no opportunity to observe category for each item thus designated as possessing high or low reliability. The results, presented in Table 8, consistently indicate that our expectations were correct. Coefficients of .70 and above were indeed found for items that were associated with consistently observable behavior, while coefficients of .30 and below were found for items that were associated with behavior that simply did not emerge and therefore could not be rated. Interestingly, items with coefficients of .30 and below (n=5) were subsumed in the independence and dependence categories, but not in the pseudoindependence category--the latter apparently reflecting behaviors that are salient, observable, ratable, and moderately reliable in terms of observer agreement.

The next step was to eliminate from the pool of checklist items those with low inter-rater reliability. We chose a cut-off of coefficients of .30 or less for this procedure. Accordingly, all items not reaching the criterion of a greater than .30 coefficient of reliability were eliminated from the checklist. By this criterion, eliminated at this stage were the following five checklist items (note--no checklist item from the pseudoindependence category had to be eliminated on the basis of this criterion):

Independence Item 14: takes appropriate initiative in problem-solving situations (r=.18)

Item 21: enters readily into new situations; is venturesome, inquisitive, etc., when appropriate to ongoing activities (r=.02)

Dependence Item 12: imitates others rather than initiating own activities (r=.12)

Item 13: gets upset when called upon in class; appears fearful, tense, timid (r=.11)

Item 26: approaches new tasks timidly and without confidence; shrinks from trying new things (r=.15)

Checklist--Item-Total Correlations (Within Each Category)

Scores for retained items were next correlated with total category scores, using the Pearson correlation method. (Nunnally [1967] has recommended that if multipoint items are employed, the regular Pearson product-moment correlation method is appropriate.)

Table 9 presents the resulting series of item-total correlation coefficients between a particular item and the total for the particular category in which it is subsumed. From this table, it can be seen that several of the items (none in the *pseudoindependence* category) fail to reach the criterion of significance of *p* is .05 or less for the particular coefficient. These items are:

Independence Item 4: volunteers contributions to class discussions and projects

Dependence Item 15: seeks to be near others

Item 27: tends to choose habitual and familiar activities

The remaining items achieve high levels of item-total correlation. These items comprise a final pool which constitutes the new, revised checklist. This new checklist, then, presented as Appendix B, contains only those items which met the criteria of sufficiently high inter-rater reliability and high item-total correlation coefficients to warrant their continued inclusion in the final pool. The revised checklist (consisting now of six items in the independence scale, four items in the dependence scale, and nine /the original group/ items in the pseudoindependence scale) is being used in a continuation of the current study to explore correlates and construct validity of the behaviors which we are able to measure reliably. This includes explorations of such variables as creativity, sociometric status, and self-concept.

Quantitative, Continuous Scores for Checklist--Z-Score Classification

It became apparent as we were working with the checklist material that it would be desirable to devise scoring procedures wherein quantitative continuous scores (within each category) could be calculated for each observed S. This would permit us, for current as well as future purposes, to be more flexible with regard to using various data-analysis procedures, for example, in ranking subjects within each of the categories, or in working with extreme subjects (determined quantitatively, rather than nominally) within a particular category.

After considering several possibilities, the procedure selected for the "scoring system" for the checklist was based on computing Z-scores (separately for each school) for each S for each item within the checklist categories (independence, dependence, pseudoindependence) based on the scoring system outlined in detail in an earlier section of this chapter. Note, Z-scores were based only on items retained after the foregoing reliability and item-analysis procedures were completed. Further, since two scores for each item per S were available (there were two observers), the procedure for obtaining a single Z-score involved computing the mean of the two observer's item score as described in the section noted above.

Within each school Z-scores were calculated per item for each S. Mean Z-scores were next computed for items within each category for each S. Table 10 presents the three mean Z-scores for each S, along with his rank in each category.[7] Table 11 presents the Ss' ranks in all three categories with Ss ordered according to rank on the independence mean Z-scores from high to low.

A "rule-of-thumb" was then devised for designating subjects for purposes of data analysis on the basis of their ranks on the three mean Z-scores as presented in Table 11. Subjects with ranks of 1 through 20 on the independence items (above median rank) are to be regarded as "independents," provided that their ranks on both of the other categories fall below their respective medians, that is, fall between 21 and 40. Subjects with ranks in the independence category of 21 through 40 are to be regarded as "nonindependents," provided that their ranks for one or the other or both of the two remaining categories, dependence or pseudoindependence fall at or above the median of 20 (i.e., 1-20). On the basis of these criteria, the designations of the Ss as presented in Table 11 resulted in classifying 13 Ss as "independent," 18 as "nonindependent," and 9 as "mixed" types.

---

[7] The data in Table 10 are presented separately according to school. Were there only chance differences in ranks due to schools, then mean of ranks for designation within each school would be 20 or as close to 20 as is possible. In actual fact, these means do indicate this to be the case. That is, as expected, there seem to be little or no school differences in assigned rank within each designation. Specifically, these means are:

School	Designation	Mean of Ranks
P.S. 90	Independence	20.50
	Dependence	20.59
(N=22)	Pseudoindependence	20.86
P.S. 175	Independence	20.50
	Dependence	20.33
(N=18)	Pseudoindependence	20.06

Reliability of "Global" Ratings Based on Checklist--Two Independent Raters

Simultaneously with, and independently of, the statistical determination of observer and item reliabilities, overall ratings for each child were made by two additional, independent raters (the Principal and Co-Principal Investigators), using the scale presented below. These raters based their ratings on two summary checklist sheets for each child (one for each independent classroom observer) which contained total frequencies for each scale point for each item on the behavior checklist. As noted, it was not possible to use the "narratives" (comments by the observers on the child's classroom behavior with regard to anything salient that might have occurred) in this process, since they did not contain sufficiently complete or consistent information from child to child to warrant their use.

Rater________________________ IDS Code No.________

Check one only:

I--Independent Highly______ Moderately______
P--Pseudoindependent Highly______ Moderately______
D--Dependent Highly______ Moderately______
X--Can't rate______

Degree of confidence in rating--check one:

3______high 2______medium 1______low

After these ratings were completed, inter-rater reliability was determined on the basis of Cohen's recommended use of weighted kappa (1968) as a coefficient of agreement for nominal scales. This measure not only corrects for chance agreement, but also allows for differential weighting of disagreements according to the degree of gravity of such disagreements. It is therefore possible to weight scale-point disagreements (i.e., highly, moderately) less heavily than categorical disagreements (i.e., independent, pseudoindependent, dependent). Table 12 contains the matrix of the weights we employed for this procedure. These weights range from 0 to 4. The larger the weight, the greater the disagreement; 0 indicates no disagreement, and 4 indicates maximum disagreement.

Table 13 shows the distribution of the global ratings made by the independent raters. These data yielded a reliability coefficient of .76 when weighted kappa was computed, taking into account scale points as well as category. This results in $\underline{z}=5.43$ ($p < .0002$, two-tailed). (Unweighted kappa, based on categorical designation only, yielded a reliability coefficient of .873, resulting in $\underline{z}=4.48$ /$\underline{p} < .0002$, two-tailed/.) It can therefore be concluded that the raters showed excellent agreement in designating the Ss into the three relevant classifications on the basis of the summary checklist data.

From Table 13, it can be seen that global ratings based on checklist summary information tend to yield little disagreement. It should also be noted that both raters were able to execute the ratings quickly and with considerable confidence. By and large, even ratings made with low confidence tended to agree with each other. The raters were under the general impression that the degree of confidence each rater tended to assign to her ratings was a question of personal "style"--for there seemed to be differences between the raters in the frequency with which "low confidence" was checked.

Reliability of Ratings--Transfer Sessions

Chapter 2 describes the manner in which ratings of behavior in the transfer sessions were made. Inter-rater agreement (for the scale just presented) was indexed by means of Cohen's weighted kappa (1968). As already noted, this computation not only corrects for chance agreement, but also allows for differential weighting of disagreements according to the degree of gravity of such disagreements (see Table 12).

Table 14 shows the frequency distribution of the ratings by paired observers of the transfer sessions. These data yielded a reliability coefficient (kappa) of .257, which, in terms of the normal curve deviate, results in $z=2.18$ ($p < .03$, two-tailed). This coefficient is based on the 46 cases remaining after the elimination of ratings in which observers expressed low confidence ($N=24$) and those designated "can't rate" ($N=6$).[9]

Table 15 summarizes the number of agreements and disagreements within and across categories. As can be seen, most of the disagreements were in regard to the independence--dependence ratings, and among these, most were in regard to the moderately independent vs. moderately dependent choices.

It is thus seen that observers' agreement in the transfer sessions, although moderately good, does not fall on the same level as that obtained for a different set of raters who made judgments on the basis of the checklist. It is highly probable that the kinds of operations involved in making such ratings differ considerably in each situation, with judgments in the transfer situations more liable to various types of error than those based on checklist information.

---

[9] Inclusion of "low confidence" ratings in the computation of weighted kappa resulted in a rather low reliability estimate prompting us to exclude those cases from the analysis described. When these cases are included (total $N=70$, for there were 6 "can't rate" designations), the findings were: weighted kappa was .182, which in terms of the normal curve deviate, results in $z=1.89$ ($p < .06$, two-tailed).

Checklist--The Relationship of Z-score Designations and "Global" Ratings

When ratings based on qualitative examination of the checklist were completed, they were compared to the designations of the experimental Ss based on their Z-score ranks for each of the categories on the checklist. Note, for this analysis, those Z-score classifications which were labeled "mixed" and those global ratings which represented disagreements between the raters or which involved a "can't rate" designation were notused. This analysis then, was based on a group of 29 Ss.

Table 16 presents the contingency classification of designations derived from each of the above sources. A chi-square analysis as to whether these two methods of classifying the Ss are associated yielded a probability level of $<.025$ and a contingency coefficient of .486. It is thus concluded that there is a positive association, established with some degree of confidence, between the method of rating behavior globally (on the basis of qualitative examination of the checklist material) and that of objectively classifying behavior on the basis of the Ss' Z-scores.

II

Data Analyses Pertaining to the Testing of Hypotheses

The current section describes findings pertaining to the following analyses: (a) global checklist ratings vs. ratings (experimental subjects only) in the transfer situation; (b) Z-score checklist designations vs. ratings in the transfer situation; and (c) distribution of experimental Ss according to "filler" status vs. global checklist ratings, Z-score checklist designations, and behavioral (transfer session) ratings. The foregoing is Part A of this section, and relates to explorations involving the experimental

Ss only. Part B of the current section pertains to experimental vs. control differences in terms of the ratings of independence or dependence made in the behavioral sessions. Analyses of results were made for schools separately and combined as well as for groups with and without Ss for whom ratings with "low confidence" were made.

A. Experimental Ss Only

Global Checklist Ratings vs. Ratings in the Transfer Situation

The next two sections present data analyses for only those Ss (experimental Ss) observed in the classrooms. These sections are addressed to the question regarding "transfer"--that is, whether designations based on classroom behavior are in any way associated with designations based on behavior in the nonclassroom situation. The first of these analyses is concerned with whether the global ratings based on the checklist are associated with status as rated in a later transfer (behavioral) situation. The findings, as presented in Table 17, indicate a lack of association between these two methods of classifying the Ss.

Z-Score Checklist Designations vs. Ratings in the Transfer Situation

Table 18 presents data pertaining to the relationship between objective Z-score classifications based on the checklist and ratings made subsequently in a nonclassroom behavioral situation. These findings also indicate a lack of association between these two methods of classifying the Ss.

The foregoing analyses (see Tables 17 and 18) do indeed suggest that insofar as the current research design in concerned, behavior as classified in the classroom is not similarly classified in subsequent nonclassroom behavioral situations. Thus, it would be reasonable to suggest that such behavior does not "transfer." The situations are so different from each other, however, that one can raise a reasonable doubt as to whether transfer effects were actually being measured by the behavioral sessions. The next chapter briefly touches on this issue again.

"Filler" Status and Three Different Ratings of Experimental Ss

As will be recalled (see Table 1), the experimental children were classified according to "filler" status, depending on the number of years of membership in the Institute classes. One of the hypotheses of this study pertains to the possibility that length of exposure to our demonstration program would be associated with status (independence, pseudoindependence, or dependence) as determined through checklist ratings, Z-score classifications, and behavioral session ratings. (See Table 19 for these frequency distributions.) Table 20 presents results of analyses pertaining to these expectations. All analyses yielded nonsignificant results, it should be noted.

Specifically, all Ss were allocated to one of five groups defined by the year the child entered the program. For example, if an S entered the Institute's program in 1964 at the prekindergarten level, he received four full years of exposure to an enriched classroom atmosphere and curriculum; if he entered the program in 1968 (an F3 subject), however, he virtually had no exposure to the Institute's program when the current study was initiated. In this sense, then, F3 subjects are similar to control subjects, at least in the beginning of the fall term of 1968. To explore the association between various measures and status of subjects in terms of length of exposure to the Institute's program, we grouped all Ss with "filler" classifications of: the full four years of exposure, through three, two, and one year of exposure (E, FK, F1, and F2 Ss) and examined their ratings in comparison to the F3 group. That is, Ss receiving one year or more of our program were compared to Ss who had entered the program the year of our investigation in terms of distributions of ratings and Z-score designations. As presented in Tables 19 and 20, data analyses indicate that there is no association between "filler" status and global checklist ratings, Z-score classifications, or behavioral session ratings.

B. Experimental vs. Control Ss (Behavioral Sessions)

Ratings of Ss in the behavioral sessions were examined in terms of experimental vs. control status of the subjects. Table 21 presents the results of chi-square analyses for subjects in both schools combined with (a) "low confidence" ratings included and (b) "low confidence" ratings eliminated. Table 21 also presents results for both schools held separately, with ratings of "low confidence" included with these analyses because of the small Ns involved. All analyses were based on four-fold tables (one degree of freedom) with Yates-corrected values of chi-square, and all comparisons involving the experimental group were based only on those Ss with one or more years of exposure to the Institute program.

For all comparisons, the N of the experimental sample was reduced from 39 to 37, because 2 Ss from Public School 175 were randomly eliminated to equate the Ns of the experimental and control samples. (These Ss, as it turned out, were also rater-disagreement Ss.) In comparisons (1) and (2), 9 F3 Ss and 13 rater-disagreement Ss were eliminated from the experimental group; in addition, 5 "low confidence" rating Ss were eliminated from the experimental group in comparison (2).

There are 37 control Ss in comparisons (1) and (2) of which 18 Ss were eliminated because they represented rater-disagreements. In addition, 4 "low confidence" control Ss were eliminated in comparison (2).

In comparison (3) there were 19 experimental and 19 control Ss

(Public School 175). The 9 Ss who were rater-disagreements and 3 F3 Ss were eliminated from the experimental group, leaving a total of 7 experimental Ss. From the control group, 11 rater-disagreements were eliminated, leaving a total of 8 control Ss.

In comparison (4), there were 18 experimental and 18 control Ss (Public School 90). Of the 18 experimental Ss, 4 rater-disagreements and 6 F3 Ss were eliminated leaving a total of 8 Ss. Of the 18 control Ss, 7 rater-disagreements were eliminated, leaving a total of 11 Ss.

All findings reported in Table 21 are nonsignificant--that is, there is no apparent association between experimental and control subject status and the ratings assigned to them based on behavior in the transfer sessions.

Chapter 4

Summary and Conclusions; Abstract of Continuation Research

This chapter will briefly recapitulate the major findings and relate these to the expectations and hypotheses described in Chapter 1. It might be noted that since the currently described research has continued into the present, and is still ongoing, major conclusions and discussion will be deferred until the present year's work is completed. This year's work (1969-1970) involves cross-validation of the checklist with a similar sample, development of reliable measures of creativity, self-concept, and sociometric choice, and exploration of the relationship of scores from the latter three areas as well as teacher-ratings of independence, to more objective measures of independence derived from the checklist. The end of this chapter presents a brief abstract of the current continuation research. It should be noted that this work is well under way.

A basic, overall expectation was that we could develop a reliable, usable classroom behavior checklist. This expectation has been met. Inter-rater reliability coefficients (see Table 7) were utilized to eliminate items of low reliability; in addition, item-total correlations for checklist items within each category (see Table 9) were further examined to eliminate additional items. The resulting checklist, presented in Appendix B, thus contains only those items which met the criteria of sufficiently high inter-rater reliability and high item-total correlation coefficients to warrant their continued inclusion in the final form of the checklist. The revised checklist is being used in a continuation of the currently described study, as noted, to explore correlates and construct validity of independent behavior.

The preceding chapter described the development of an objective Z-score based on checklist findings as well as a systematic "rule" for classifying subjects based on their Z-score profiles. These designations are presented in Table 11, and resulted in classifying 13 Ss as "independent," 18 as "nonindependent," and 9 as "mixed" types. (The latter, it should be noted, is not synonymous with the "pseudoindependent" category.) Global raters (who were not in the classroom) independently rated the Ss on the basis of summary checklist data (see Table 13) with an extremely high degree of reliability--that is, with little or no disagreement.

Table 16 presents data pertaining to the association between global ratings and Z-score designations of the Ss. It was concluded that there is a positive association, established with some degree of confidence, between these checklist methods of categorizing behavior.

Ratings of observers in the transfer situation (see Table 14) were less reliable, it should be noted, than those based on summary checklist data. Several possibilities can be offered to "explain" the latter finding: (a) four different raters (rather than two, as planned) were used in the transfer session; (b) ratings in the transfer session were based on one observation session, whereas data obtained in the classroom were obtained over many sessions of observation; (c) the tasks and behavioral situations we utilized were in many ways unique to the observers and to the subjects alike; despite training, the raters were perhaps not comfortable in assigning ratings to the Ss based on such an unconventional approach; and (d) the nature of the behavioral session was such that it might simply not have elicited reliably ratable behavior (at least to the same extent as did the classroom situation).

For the experimental Ss, overall findings indicate that the designations based on classroom behavior are not associated with designations derived from the transfer situation. Thus, Tables 17 and 18 indicate that insofar as the current research is concerned, behavior classified in the classroom is not "carried over" to nonclassroom behavioral situations. Based on classroom data alone, however (see Table 13), experimental subjects were judged more frequently by independent raters to be independent than otherwise; further, as this table also shows, such raters used the category of pseudoindependence even more frequently than dependence, suggesting that the pseudoindependent category, as expected, is meaningful in terms of characterizing classroom behavior. Also, classroom observers were particularly reliable in judging pseudoindependent behavior; and item-total Pearson correlation coefficients in the pseudoindependent category were all significant.

As presented in Tables 19 and 20, data analyses indicate that there is no association between "filler" status and the global checklist ratings, Z-score classifications, and behavioral session ratings of the experimental subjects. Finally, Table 21 indicates that there is no apparent association between experimental and control subject-status and the ratings assigned to them based on behavior in the behavioral sessions.

Overall, the foregoing findings indicate lack of support for most of the hypotheses and expectations as presented in Chapter 1, with the exception of those pertaining to the checklist, and its development and use. Indeed, it does seem that the "transfer" session was an inadequate method for measuring "transfer" effects, and that although we are seemingly eliciting independent behavior in the classroom, the experimental Ss, when compared to experimental F3 "filler" Ss and to the control group Ss, did not appear to differ in terms of their assigned ratings with any degree of statistical confidence. Some reasons for greater unreliability of the ratings in these situations were presented earlier. It may also be that the significant and crucial variable is the striking differences between the classroom situation and the behavioral situation. In some respects, the behavioral situation resembled the typical control classroom more than did the experimental classroom. In this sense, Institute independent children might well have not "carried" their independent behavior over into the new situation. Conversely, control S̲s might have felt freer to show independent behavior in a situation which was structured to elicit such behavior more than did their typical classroom situation.

Although a complete discussion of these points will be presented in the 1970 Final Report, it should be noted here that the literature presents some evidence that dependency may not be a unitary trait but rather is a rubric for different behaviors. Thus, a child would be rated differently depending on the aspect of dependency measured. Furthermore, some studies indicate that the situational aspects in which the dependency is measured are of importance. It is likely that this evidence is relevant to the study of independence, too.

In any event, we may be forced to conclude that independence does not "carry over" into other situations--at least in this age group--regardless of that situation. Further, the nature of the task, situation, and the requirement placed before the child, seem to be of key significance. Questions of construct validity, that is, just what independence is and what its correlates are seem to be crucial, at this point of our studies, and it is to these issues that our current work is addressed. We are content that we have developed a usable instrument for doing so, at least in terms of classroom activity. Further, since many "conventional" classrooms (kindergarten, for example, or prekindergarten) provide the kind of milieu

8

that allows observable behavior to emerge (of the sort our checklist is designed to measure), we are hopeful that our cross-validated checklist can be useful at younger age-levels.

Abstract of Continuation Research (1969-1970)

(1) Objectives. In our original study, we proposed to assess the degree to which the Institute's ongoing enrichment program in several Harlem public schools has succeeded in fostering and encouraging independent, autonomous behavior among its pupils. This evaluation has been conducted through the use of behavioral, rather than standardized test criteria, both in the classroom and in a "transfer" situation, to explore whether or not the changes that occur in the classroom do emerge outside of that "programed" setting. This investigation, and the proposed continuation research, have implications not only for scholastic achievement but also for the child's overall adjustment in daily life. If we are indeed training for, and reinforcing independent behavior in Institute children, we must be assured that independence should be an educational goal. For the above reason, we proposed to continue our current research to determine what associated, correlative behaviors and characteristics accompany independent, pseudo-independent, or dependent behavior. The specific objectives of this continuation research are the following:

(a) To replicate and cross-validate the revised classroom behavior checklist with a new but similar pupil sample. This replication would hopefully enable us to provide the professional community with a reliable, cross-validated instrument for the measurement of independent, pseudoindependent, and dependent behavior.

(b) To explore the relationship of teachers' ratings of children as to independent, pseudoindependent, and dependent behavior to behavioral checklist scores based on observations of these same children.

(c) To explore the sociometric concomitants of independence, pseudoindependence, and dependence.

(d) To explore the relationship of creativity measures and scores to designations of independence, pseudoindependence, and dependence derived from the checklist, and to develop new methods for assessing the creativity dimensions in this sample.

(e) To explore the relationship of self-concept (as measured by Q-methodology) to designations of independence, pseudoindependence, and dependence derived from the checklist.

(f) To examine a number of additional, subsidiary relationships that can shed light on the behaviors relevant to the area of our investigation.

(2) Procedures. Our sample is composed of all children in as many 1969-1970 third-grade Institute demonstration classrooms as are available. Third-graders have been chosen as these children have had the longest exposure to the Institute program and should, therefore, show the greatest effects of this intervention. Within our sample will be "fillers," those children who have been added to the program because of attrition in the initial group. The presence of "fillers" will enable us to determine whether behavioral change is a function of the length of time the child has spent in the program. Using the revised behavior checklist, we plan to proceed as follows:

(a) Early in the school year, two observers will independently record, by frequency tallies on the checklist, each child's behavior during five-minute periods over a total of 20 sessions on different days (this step is currently near completion).

(b) An initial set of sociometric measures will be obtained in each classroom by personnel other than those making the behavioral observations.

This step has been completed. These same measures will be collected again at the end of the year.

(c) Teachers' ratings of the same children in regard to independence, pseudoindependence, and dependence will be obtained at both the beginning and the end of the school year. The first wave of teachers' ratings has been obtained.

(d) Measures of creativity will be developed and administered to the Ss in the sample.

(e) Q-sort techniques will be developed and administered to all of the Ss in the samples simultaneously with step (d) to measure self-concepts and self-perceptions.

At each step along the way, it might be noted, stress will be placed on developing reliable and objective instruments and measures.

References

Asch, S.E. Studies of independence and conformity: I. a minority of one against a unanimous majority. *Psychological Monographs*, 1956, 70, No. 9 (Whole No. 416).

Beller, E.K. Dependency and independency in young children. *Journal of Genetic Psychology*, 1955, 87, 23-35.

Cervenka, E.J. The measurement of bilingualism and bicultural socialization of the child in the school setting: the development of instruments. Final report on Headstart Evaluation and Research 1966-67 (contract #66-1).

Cohen, J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 1968, 70, 213-220.

Crandall, V.C. Achievement behavior in young children. In W. Hartup and L. Smothergill (Eds.), *The young child*. Washington: NAEYC, 1967.

Davidson, H.H., & Greenberg, J.W. *Traits of school achievers from a deprived background*. New York: Simon and Schuster, 1967.

Deutsch, C., and Schumer, F. *Brain-damaged children: a modality-oriented exploration of performance*. New York: Brunner/Mazel, in press.

Edwards, A. L. *Statistical methods for the behavioral sciences*. New York: Holt, Rinehart, and Winston, 1954.

Eells, K., Davis, A., Havighurst, R.J., Herrick, V.E., & Tyler, R.W. *Intelligence and cultural differences: a study of cultural learning and problem-solving*. Chicago: University of Chicago Press, 1951.

Finney, J.C. Some maternal influences on children's personality and character. *Genetic Psychology Monographs*, 1961, 63, 199-278.

Gewirtz, J.L. A factor analysis of some attention-seeking behaviors in young children. Child Development, 1956, 27, 17-36.

Guidelines for testing minority group children. Journal of Social Issues, 1964, 20, 129-145.

Hartup, W.W. Dependence and independence. In H.W. Stevenson (Ed.), Child psychology: The 62nd yearbook of the National Society for the Study of Education, Part I. Chicago: The National Society for the Study of Education, 1963.

Minuchin, S., Montalvo, B., Guerney, B.G., Jr., Rosman, B.L., & Schumer, F. Families of the slums: an exploration of their structure and treatment. New York: Basic Books, Inc., 1967.

Nunnally, J. Psychometric theory. New York: McGraw-Hill, 1967.

Powledge, F. To change a child: a report on the Institute for Developmental Studies. Chicago: Quadrangle Books, 1967.

Zubin, J., Eron, L.D., & Schumer, F. An experimental approach to projective techniques. New York: John Wiley and Sons, Inc., 1965.

Table 1

Experimental Sample Observed in the Classrooms: Mean Age and Sex for Subjects in Public Schools 90 and 175 Classified According to Years of Exposure to the Enrichment Program

Entered IDS Classes	N	Sex	Subject Classification[a]	Mean Age (Months)[b]
		Public School 90		
Pre-K, 1964-1965 (4 years of exposure)	4	M--4 F--0	E	97.25
Kindergarten, 1965-1966 (3 years of exposure)	2	M--1 F--1	FK	101.00
First Grade, 1966-1967 (2 years of exposure)	6	M--3 F--3	F1	99.50
Second Grade, 1967-1968 (1 Year of exposure)	1	M--0 F--1	F2	103.00
Third Grade, 1968-1969 (0 years of exposure)	6	M--3 F--3	F3	98.17
Total, P.S. 90	19	M--11 F--8		98.84
		Public School 175		
Pre-K, 1964-1965 (4 years of exposure)	14	M--8 F--6	E	98.35
First Grade, 1966-1967 (2 years of exposure)	2	M--1 F--1	F1	99.50
Second Grade, 1967-1968 (1 year of exposure)	2	M--2 F--0	F2	99.00
Third Grade, 1968-1969 (0 years of exposure)	3	M--2 F--1	F3	98.00
Total, P.S. 175	21	M--13 F--8		98.48
Total, Schools Combined	40	M--24 F--16	18 Es 2 FKs 8 F1s 3 F2s 9 F3s	98.68

Note--There were no FK subjects in Public School 175 in the third grade.

[a] F categories contain "filler" Ss, added to the experimental group through the years because of attrition.

[b] As of September, 1968. Mean age for all Ss is just under 8 years, 3 months.

Table 2

Experimental Sample Observed in the Classrooms: Mean Age for Subjects Classified According to Years of Exposure to the Enrichment Program for Both Schools Combined

Subject Classification	N	Mean Age[a] (Months)
E	18	98.11
FK	2	101.00
F1	8	99.50
F2	3	100.33
F3	9	97.89
Total	40	98.68

[a] As of September, 1968.

Table 3

Behavioral Transfer Sessions--Experimental Subjects: Mean Age and Sex for Subjects in Public Schools 90 and 175 Classified According to Years of Exposure to the Enrichment Program

Entered IDS Classes	N	Sex	Subject Classification[a]	Mean Age (Months)[b]
		Public School 90		
Pre-K, 1964-1965 (4 years of exposure)	4	M--4 F--0	E	97.25
Kindergarten, 1965-1966 (3 years of exposure)	2	M--1 F--1	FK	101.00
First Grade, 1966-1967 (2 years of exposure)	6	M--3 F--3	F1	99.50
Second Grade, 1967-1968 (1 year of exposure)	1	M--0 F--1	F2	103.00
Third Grade, 1968-1969 (0 years of exposure)	5	M--3 F--2	F3	98.40
Total, P.S. 90	18	M--11 F--7		99.06
		Public School 175		
Pre-K, 1964-1965 (4 years of exposure)	14	M--8 F--6	E	98.35
First Grade, 1966-1967 (2 years of exposure)	2	M--1 F--1	F1	99.50
Second Grade, 1967-1968 (1 year of exposure)	2	M--2 F--0	F2	99.00
Third Grade, 1968-1969 (0 years of exposure)	3	M--2 F--1	F3	98.00
Total, P.S. 175	21	M--13 F--8		98.48
Total, Schools Combined	39	M--24 F--15	18 Es 2 FKs 8 F1s 3 F2s 8 F3s	98.76

Note--There were no FK subjects in P.S. 175 in the third grade.

[a] F categories contain "filler" Ss, added to the experimental group through the years because of attrition.

[b] As of September, 1968. Mean age for all Ss is just under 8 years, 3 months.

Table 4

Experimental Sample in the Behavioral Transfer Sessions: Mean Age for Subjects Classified According to Years of Exposure to the Enrichment Program for Both Schools Combined

Subject Classification	N	Mean Age[a] (Months)
E	18	98.11
FK	2	101.00
F1	8	99.50
F2	3	100.33
F3	9	98.25
Total	39	98.76

[a] As of September, 1968.

Table 5

Behavioral Transfer Sessions--Control Subjects: Mean Age and Sex for Subjects in Public Schools 90 and 175

	N	Sex	Mean Age (Months)[a]
		Public School 90[b]	
	18	M--7 F--11	99.47[b]
		Public School 175	
	19	M--8 F--11	97.84
Total	37	M--15 F--22	98.61[b]

Note--Although 39 Ss were run in the transfer sessions, only 37 were usable for statistical analysis because of certain errors in sample selection. The Ss eliminated were both from Public School 175. To compensate for this reduction in sample size, when experimental and control differences are explored, two Ss from Public School 175 were randomly eliminated from the experimental sample to maintain the initial equality in sample size for the two treatment groups. In statistical analyses involving only the experimental sample, all 39 experimental Ss were used.

[a] As of September, 1968.

[b] Total n from Public School 90 on which mean age is based is 17 subjects because of incomplete school data at this school for one S. Mean age for control Ss is therefore based on 36 cases, and is just under 8 years, 3 months, making control and experimental subjects comparable in this regard (see Table 1).

Table 6

Distribution of Subjects by School and Sex in the Transfer Sessions[a]

		Experimental		Control		
Session		M	F	M	F	N
		Public School 90				
1		2	1	1	2	6
2		1	2	1	2	6
3		1	2	1	2	6
4		3	0	1	2	6
5		1	1	1	1	4
6		2	0	1	1	4
7		1	1	1	1	4
	N	11	7	7	11	36
		Public School 175				
1		3	0	1	2	6
2		1	2	1	2	6
3		3	0	1	2	6
4		2	0	2	0	4
5		1	2	1	2	6
6		1	2	3	0	6
7		1	1	0	2	4
8		1	1	1	1	4
	N	13	8	10	11	42
	Total	24	15	17	22	78

[a] See Note, Table 5.

Table 7

Inter-rater Reliability Coefficients of Behavior Checklist Items

Item		Product-Moment Correlation Coefficients[a]
	Independence	
(4)	volunteers contributions to class discussions and projects	.31
(5)	helps, sympathizes, shows consideration of others; is thoughtful; offers praise, guidance, assurance	.35
(6)	sits quietly and attentively when task requires	.71
(9)	utilizes available free time constructively	.50
(11)	withstands interference while engaged in individual activities, e.g., doing difficult assignment, a puzzle, painting	.40
(14)	takes appropriate initiative in problem-solving situations	.18
(20)	tries to figure out things for himself before calling on teacher (leader) or other children for help	.49
(21)	enters readily into new situations; is venturesome, inquisitive, etc., when appropriate to ongoing activities	.02
(25)	attempts to resolve difficulties that arise between himself and other children without appealing to teacher (leader)	.50
	Dependence	
(1)	withdraws in the face of difficulty	.38
(12)	imitates others rather than initiating own activities	.12
(13)	gets upset when called upon in class; appears fearful, tense, timid	.11
(15)	seeks to be near others	.64
(17)	requires close assistance or direction in order to work at a task	.70
(19)	lets other children impose on him or boss him around	.45
(23)	seeks constant recognition and/or approval	.69
(26)	approaches new tasks timidly and without confidence; shrinks from trying new things	.15
(27)	tends to choose habitual and familiar activities	.68
	Pseudoindependence	
(2)	is inattentive; is easily distracted by things going on around him	.54
(3)	uses materials and equipment in a rough or destructive manner	.58
(7)	brags, shows off, displays exaggerated opinion of own abilities	.58
(8)	seeks attention, i.e., in a manner that seems designed for active notice rather than stemming from a need for help	.77
(10)	interferes with the work of others	.72
(16)	does not follow directions	.60
(18)	seeks to do things differently from others even when own method is not effective	.53
(22)	does not pursue task to completion	.35
(24)	speaks out of turn; interrupts others; takes other's turn, etc.	.95

[a] For all reported coefficients, N=40.

Table 8

Frequencies of No Opportunity to Observe Category for Behavior Checklist Items with Reliability Coefficients of .30 and Below and .70 and Above

		Frequency--No Opportunity to Observe
	Coefficients of .30 and Below	
	Independence	
(14)	takes appropriate initiative in problem-solving situations	466
(21)	enters readily into new situations; is venturesome, inquisitive, etc., when appropriate to ongoing activities	633
	Dependence	
(12)	imitates others rather than initiating own activities	618
(13)	gets upset when called upon in class; appears fearful, tense, timid	586
(26)	approaches new tasks timidly and without confidence; shrinks from trying new things	601
	Coefficients of .70 and Above	
	Independence	
(6)	sits quietly and attentively when task requires	33
	Dependence	
(17)	requires close assistance or direction in order to work at a task	83
	Pseudoindependence	
(8)	seeks attention, i.e., in a manner that seems designed for active notice rather than stemming from a need for help	36
(10)	interferes with the work of others	35
(24)	speaks out of turn; interrupts others; takes other's turn, etc.	124

Table 9

Item-Total Pearson Product-Moment

Coefficients for Checklist Items[a]

Checklist Item No.	r	Corrected[b] r	p[c]
		Independence	
4	.544	.245	n.s.
5	.679	.530	.005
6	.755	.678	.005
9	.692	.579	.005
11	.640	.538	.005
20	.675	.547	.005
25	.622	.344	.025
		Dependence	
1	.700	.518	.005
15	.367	.106	n.s.
17	.804	.661	.005
19	.536	.407	.01
23	.802	.701	.005
27	.603	.262	n.s.
		Pseudoindependence	
2	.766	.678	.005
3	.691	.659	.005
7	.767	.741	.005
8	.825	.768	.005
10	.816	.735	.005
16	.919	.893	.005
18	.606	.506	.005
22	.722	.575	.005
24	.563	.377	.01

[a] Items for which inter-rater reliability coefficients were .30 or less were previously dropped from the checklist.

[b] Corrected for spuriousness (to eliminate item's own contribution to total score).

[c] Items marked n.s. had p values of .05 or higher.

Table 10

Mean Z-Scores and Ranks for Independence, Dependence, and Pseudoindependence Checklist Categories

N=40

	Independence		Dependence		Pseudoindependence	
Subject Code # P.S. 90	Mean Z-Score	Rank	Mean Z-Score	Rank	Mean Z-Score	Rank
0000	0.08	18	-0.52	16	-1.03	4
937	0.24	20	-0.28	21	1.18	36
939	1.09	36	-1.04	4	-1.31	2
943	0.67	30	-0.56	15	-0.81	7
945	-1.66	3	-0.13	22	0.31	29
947	-1.55	5	2.36	39	0.89	32
948	-1.08	8	0.72	31	0.58	31
955	-0.96	11	-0.11	23	1.39	37
957	0.39	23	0.24	27	0.29	28
958	-0.02	16	-0.72	12	-0.13	25
961	1.03	35	-1.22	2	-1.25	3
962	-1.52	6	0.72	32	0.24	27
967	-0.23	15	-0.72	14	-0.05	26
968	0.75	31	-1.05	3	-0.99	5
1004	-0.96	10	0.86	34	0.33	30
1005	0.60	28	0.20	26	-0.23	21
1010	-0.69	12	-0.52	16	1.13	35
1011	1.24	38	-1.43	1	-1.43	1
1012	1.21	37	-0.74	11	-0.68	10
703	-1.13	7	2.15	40	2.61	39
932	0.34	22	1.12	35	-0.49	17
934	1.45	40	0.66	29	-0.55	14
P.S. 175						
851	1.34	39	-0.49	18	-0.61	11
852	0.08	17	-0.87	8	-0.57	13
854	-1.95	2	1.38	37	2.69	40
855	-0.24	14	0.00	24	1.01	34
872	-1.66	4	0.75	33	2.17	38
875	0.56	26	-0.76	10	-0.55	15
878	1.01	34	-1.00	5	-0.86	6
879	-2.09	1	2.74	38	1.01	33
881	-0.35	13	-0.72	13	-0.18	22
883	-0.97	9	-0.48	19	-0.16	24
884	0.78	32	-0.91	7	-0.69	9
886	0.43	24	-0.36	20	-0.33	20
1003	0.99	33	-0.98	6	-0.59	12
1019	0.32	21	-0.79	9	-0.38	19
1020	0.61	29	0.59	28	-0.42	18
1021	0.58	27	0.05	25	-0.79	8
1023	0.11	19	1.14	36	-0.18	23
1024	0.46	25	0.70	30	-0.55	16

Table 11

Ranks of Checklist Mean Z-Scores for Independence, Dependence, and Pseudoindependence[a]

N=40

Subject Code No.	Rank			Designation[b]
	Independence	Dependence	Pseudoindependence	
934	1	12	27	M
851	2	23	30	I
1011	3	40	40	I
1012	4	30	31	I
939	5	37	39	I
961	6	39	38	I
878	7	36	35	I
1003	8	35	29	I
884	9	34	32	I
968	10	38	36	I
943	11	26	34	I
1020	12	13	23	M
1005	13	15	20	M
1021	14	16	33	M
875	15	31	26	I
1024	16	11	25	M
886	17	21	21	I
957	18	14	13	M
932	19	6	24	M
1019	20	32	22	I
937	21	20	5	NI
1023	22	4	18	NI
0000	23	24	37	M
852	24	33	28	M
958	25	29	16	NI
967	26	27	15	NI
855	27	17	7	NI
881	28	28	19	NI
1010	29	25	6	NI
955	30	18	4	NI
1004	31	7	11	NI
883	32	22	17	NI
948	33	10	10	NI
703	34	3	2	NI
962	35	9	14	NI
947	36	2	9	NI
872	37	8	3	NI
945	38	19	12	NI
854	39	5	1	NI
879	40	1	8	NI

[a] Subjects are ordered by rank on Independence scale.

[b] Designations are: I = Independent; NI = Nonindependent; and M = Mixed type.

Table 12

Weighting Matrix for Measuring Disagreement between Paired Observers

	Observer A					
	Independent		Pseudoindependent		Dependent	
Observer B	Highly	Moderately	Highly	Moderately	Highly	Moderately
Independent						
Highly	0	1	4	3	4	3
Moderately	1	0	3	2	3	2
Pseudoindependent						
Highly	4	3	0	1	4	3
Moderately	3	2	1	0	3	2
Dependent						
Highly	4	3	4	3	0	1
Moderately	3	2	3	2	1	0

Table 13

Distribution of Two Independent Raters' "Global" Ratings Based on Summary Checklist Data by Category and Degree

N=40

		Breakdown of Agreements Within Categories According to Degree		
Category	Agreements	High High	High Moderate	Moderate Moderate
Independent	29	13	8	8
Dependent	1	0	0	1
Pseudo-independent	7	0	0	7
	37			
Non-agreements or can't rate[a]	3			
Total *N*	40			

[a] For two *S*s, the raters disagreed on category; for one *S* there was a *can't rate* designation.

Table 14

Frequency Distribution of Ratings by Paired Observers of Transfer Sessions[a]

	Observer A						
	Independent		Pseudoindependent		Dependent		
Observer B	Highly	Moderately	Highly	Moderately	Highly	Moderately	Total
Independent							
Highly	0	2	0	0	1	1	4
Moderately	6	10	0	0	1	2	19
Pseudoindependent							
Highly	0	0	0	0	0	0	0
Moderately	0	0	0	0	1	0	1
Dependent							
Highly	0	0	0	0	1	3	4
Moderately	3	5	0	1	1	8	18
Total	9	17	0	1	5	14	46

[a] Includes only those ratings in which the observers expressed high or moderate confidence. Total N was 76 (39 experimental Ss and 37 control Ss). Ratings in which observers expressed low confidence (N=24) were excluded, as are those designated "can't rate" (N=6).

Table 15

Summary of Agreements and Disagreements Within and Across Categories of Ratings in Transfer Sessions[a]

Frequency of Agreements Within Categories

Independence	
Highly--Highly	0
Moderately--Moderately	10
Moderately--Highly	8
n	18

Pseudoindependence	
Highly--Highly	0
Moderately--Moderately	0
Moderately--Highly	0
n	0

Dependence	
Highly--Highly	1
Moderately--Moderately	8
Moderately--Highly	4
n	13
N	31

Frequency of Disagreements Across Categories

Pseudoindependence--Independence	
Highly pseudo.--Highly inde.	0
Moderately pseudo.--Highly inde.	0
Highly pseudo.--Moderately inde.	0
Moderately pseudo.--Moderately inde.	0
n	0

Independence--Dependence	
Highly inde.--Highly depend.	1
Moderately inde.--Moderately depend.	1
Highly inde.--Moderately depend.	4
Moderately inde.--Moderately depend.	7
n	13

Pseudoindependence--Dependence	
Highly pseudo.--Highly depend.	0
Moderately pseudo.--Highly depend.	1
Highly pseudo.--Moderately depend.	0
Moderately pseudo.--Moderately depend.	1
n	2
N	15

---

[a] See footnote, Table 14.

Table 16

Checklist Designations: Global Ratings vs. Z-Score Classifications[a]

$\underline{N}=29$[b]

Global Rating	Z-Score Classification	
	Independent	Non-Independent
Independent	13	8
Dependent	0	1
Pseudoindependent	0	7
Totals	13	16

[a] Data analysis: Chi-square = 8.976; two degrees of freedom; $\underline{p} < .025$; contingency coefficient =.486.

[b] N of observed Ss in the classrooms = 40. Eliminated from this analysis were: 9 Ss (see Table 11) whose status was designated mixed on the basis of Z-score profiles (note, one of these Ss was also a "can't rate" as to globally rated status); and 2 Ss who represented disagreements as to category between the global checklist raters.

Table 17

Global Checklist Ratings vs. Behavioral Transfer Session Ratings[a]

N=22[b]

Behavioral Transfer Session Rating	Global Checklist Rating	
	Independent	Non-Independent[c]
Independent	8	3
Non-Independent	11	0
Totals	19	3

---

[a] Data analysis: Chi-square=1.544; one degree of freedom; p value is non-significant; phi coefficient =.27. Chi-square value is Yates-corrected.

[b] N of observed Ss in the classrooms = 40; one S was absent from the transfer session so that all analyses for experimental Ss in the transfer sessions are based on an N of 39. Of these Ss, for purposes of the above analysis, the following Ss were eliminated: 14 Ss representing disagreements between raters in the behavioral sessions, and the three Ss representing disagreements or "can't rate" of the global raters of the checklist (interestingly, 2 of these 3 Ss also represented disagreements between raters in the transfer sessions).

[c] This category represents a "collapsed" category of Dependent and Pseudoindependent global ratings.

Table 18

Z-Score Checklist Designations vs. Ratings in the Transfer Situation[a]

N=18[b]

Ratings in Transfer Sessions	Z-Score Checklist Designation	
	Independent	Non-Independent
Independent	4	4
Dependent	6	3
Pseudoindependent	1	0
Totals	11	7

---

[a] Data analysis: Chi-square = 1.168; two degrees of freedom; p value is nonsignificant; contingency coefficient = .247.

[b] N of observed Ss in the classroom = 40; one S was absent from the transfer session so that all analyses for experimental Ss in the transfer sessions are based on an N of 39. Of these Ss, for purposes of the above analysis, the following Ss were eliminated: 9 Ss for whom the Checklist Z-score designations were "mixed" (four of these were also disagreements between the raters in the behavioral sessions); and 12 Ss representing disagreements between raters in the behavioral sessions.

Table 19

Frequency Distributions of Global Checklist Ratings, Z-Score Checklist Designations, and Behavioral Sessions Ratings (Experimental Subjects Only) by Length of Exposure to Institute Program ("Filler Status")

Subject Classification	N[a]	Length of Exposure to Institute Program[b]				
		E	FK	F1	F2	F3
(1)						
Global Checklist Rating	37					
Independent		11	1	7	2	8
Dependent		0	0	0	1	0
Pseudoindependent		4	1	1	0	1
(2)						
Z-Score Checklist Designation	31					
Independent		4	1	4	1	3
Non-Independent		10	1	4	1	2
(3)						
Behavioral Session Rating	23					
Independent		3	0	4	1	4
Dependent		4	1	2	1	2
Pseudoindependent		0	0	0	0	1

---

[a] See Footnote (a) Table 20 for explanation of these Ns.

[b] Groups designated as E entered the program at the pre-kindergarten level in September, 1964.
Groups designated as FK entered the program at the kindergarten level in September, 1965.
Groups designated as F1 entered the program at the first grade level in September, 1966.
Groups designated as F2 entered the program at the second grade level in September, 1967
Groups designated as F3 entered the program at the third grade level in September, 1968.

Table 20

Checklist and Transfer Session Ratings of Experimental Subjects and Length of Exposure to the Institute's Program ("Filler" Status)

	N[a]	Comparison Groups[b]	Chi-Square Value	Degrees of Freedom	p value
(1)	37	Global Checklist Ratings X F3 Ss vs. all others (3x2 table)	.872	2	ns
(2)	31	Z-score Checklist Designations X F3 Ss vs. all others (2x2 table)	.799[c]	1	ns
(3)	23	Behavioral (Transfer) Session Ratings X F3 Ss vs. all others (3x2 table)	2.847	2	ns

---

[a] N=40 (classroom observations). For comparison (1), N=37, since 3 Ss represented disagreements or can't rate in re: the global ratings. For comparison (2), N=31 since there were 9 "mixed" Z-score classifications. For comparison (3) N=23, since there were 39 experimental Ss in the transfer session and 16 represented disagreements between the raters.

[b] See Table 19 for explanations of "filler" status. For example, F3 refers to all experimental subjects who entered the Institute's program in 1968.

[c] Yates-corrected (2x2 table).

Table 21

Chi-Square Comparisons of Ratings of Behavioral Sessions: Experimental vs. Control Subjects for Both Schools Combined and for Schools Held Separately[a]

Both Schools Combined

Comparison	Rating	Exper.	Control	*N*[b]	Chi-Square	*p* value
(1) With "low confidence" ratings included	Indep.	7	13		.863	ns
	Depend.	8	6			
N		15	19	34		
(2) With "low confidence" ratings excluded (even if *one* rater used this category)	Indep.	4	9		.327	ns
	Depend.	6	6			
N		10	15	25		
(3) Public School 175 (With "low confidence" ratings included)						
	Indep.	3	5		.059	ns
	Depend.	4	3			
N		7	8	15		
(4) Public School 90 (With "low confidence" ratings included)						
	Indep.	4	8		.283	ns
	Depend.	4	3			
N		8	11	19		

[a] All comparisons are based on four-fold tables (one degree of freedom) with chi-square values that are Yates-corrected. The experimental group contains only those *S*s who have had one or more years of exposure to the Institute program.

[b] See text, Part B, end of Chapter 3, for explanation of these *N*s.

Geometric Design (A) for Task (2)

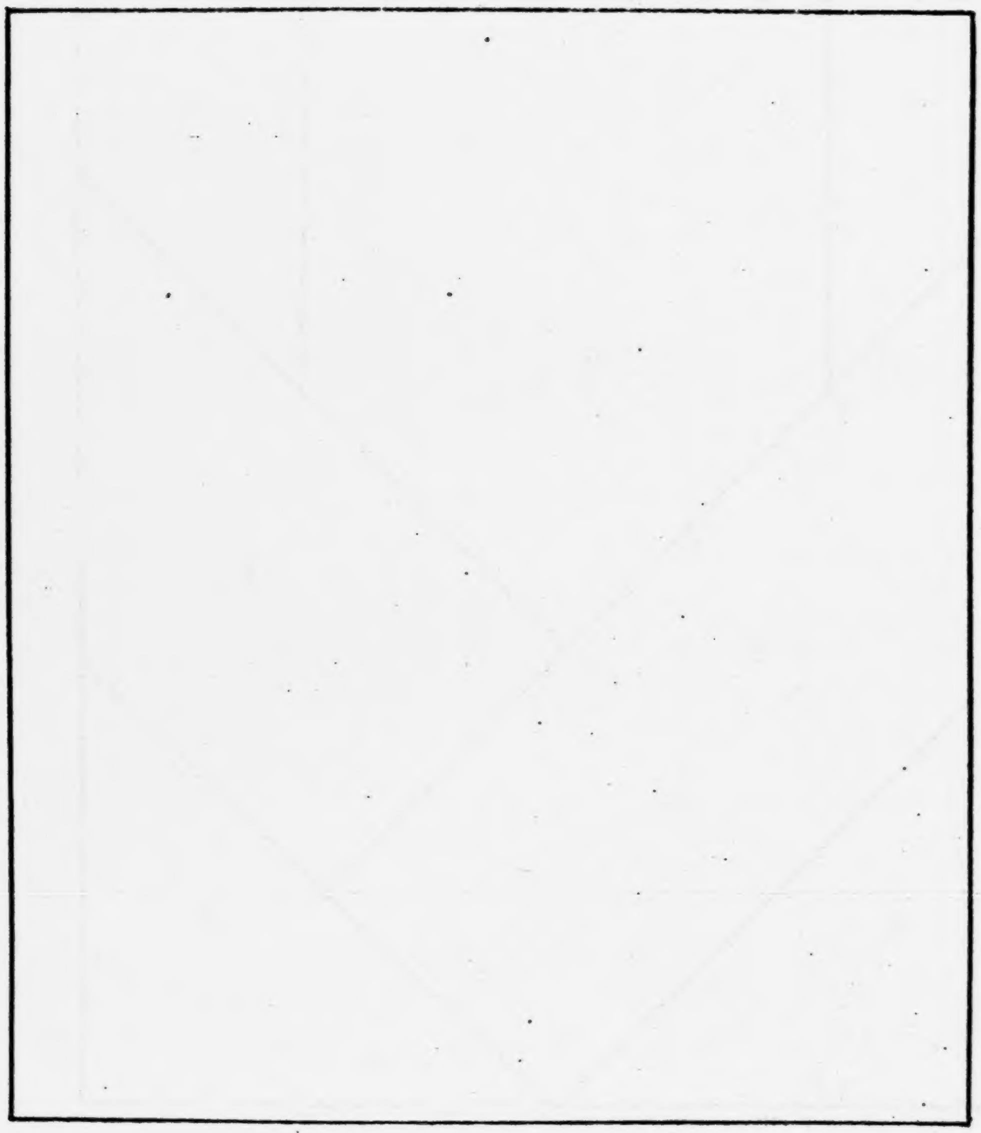


Figure 1

Solution for Design (A)--Task (2)

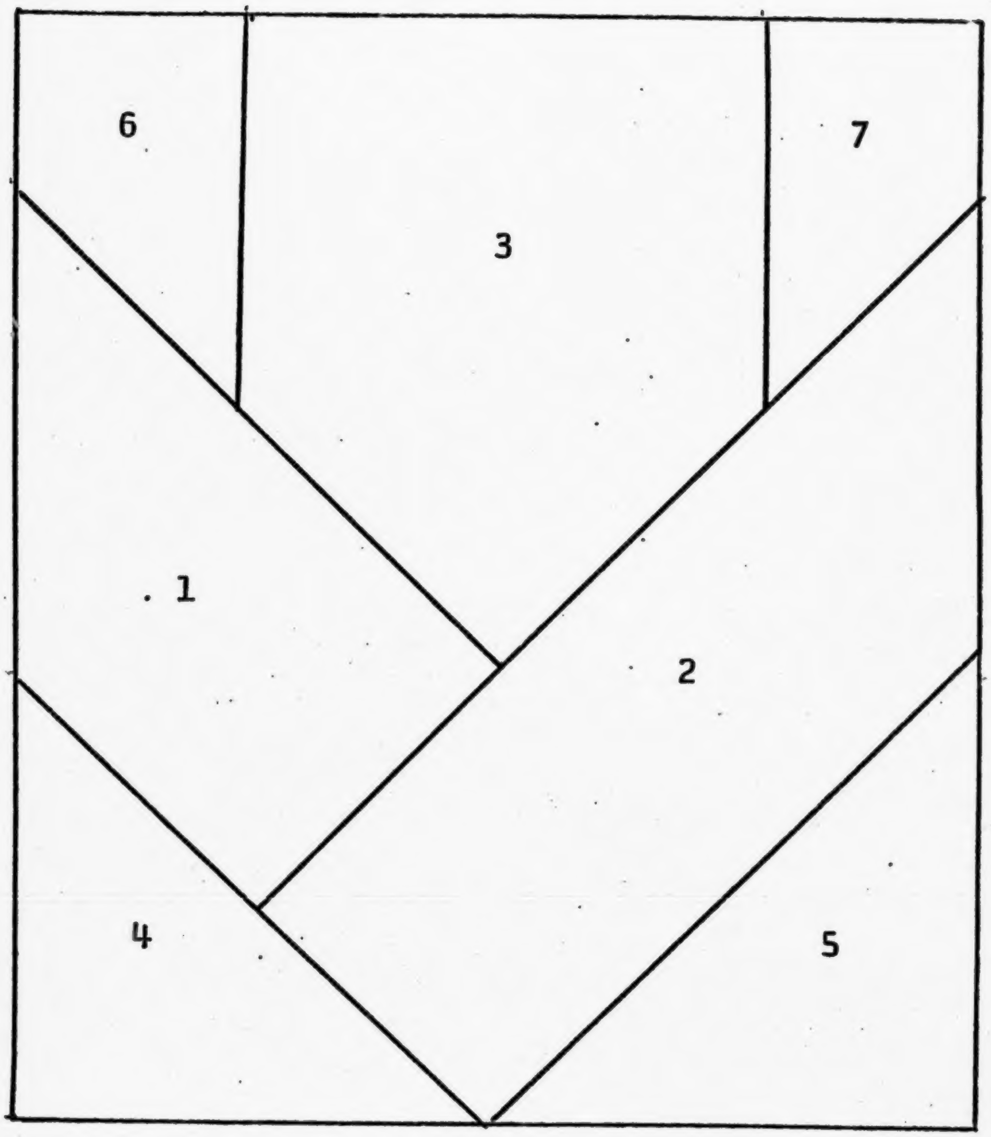


Figure 2

Geometric Design (B) for Task (2)

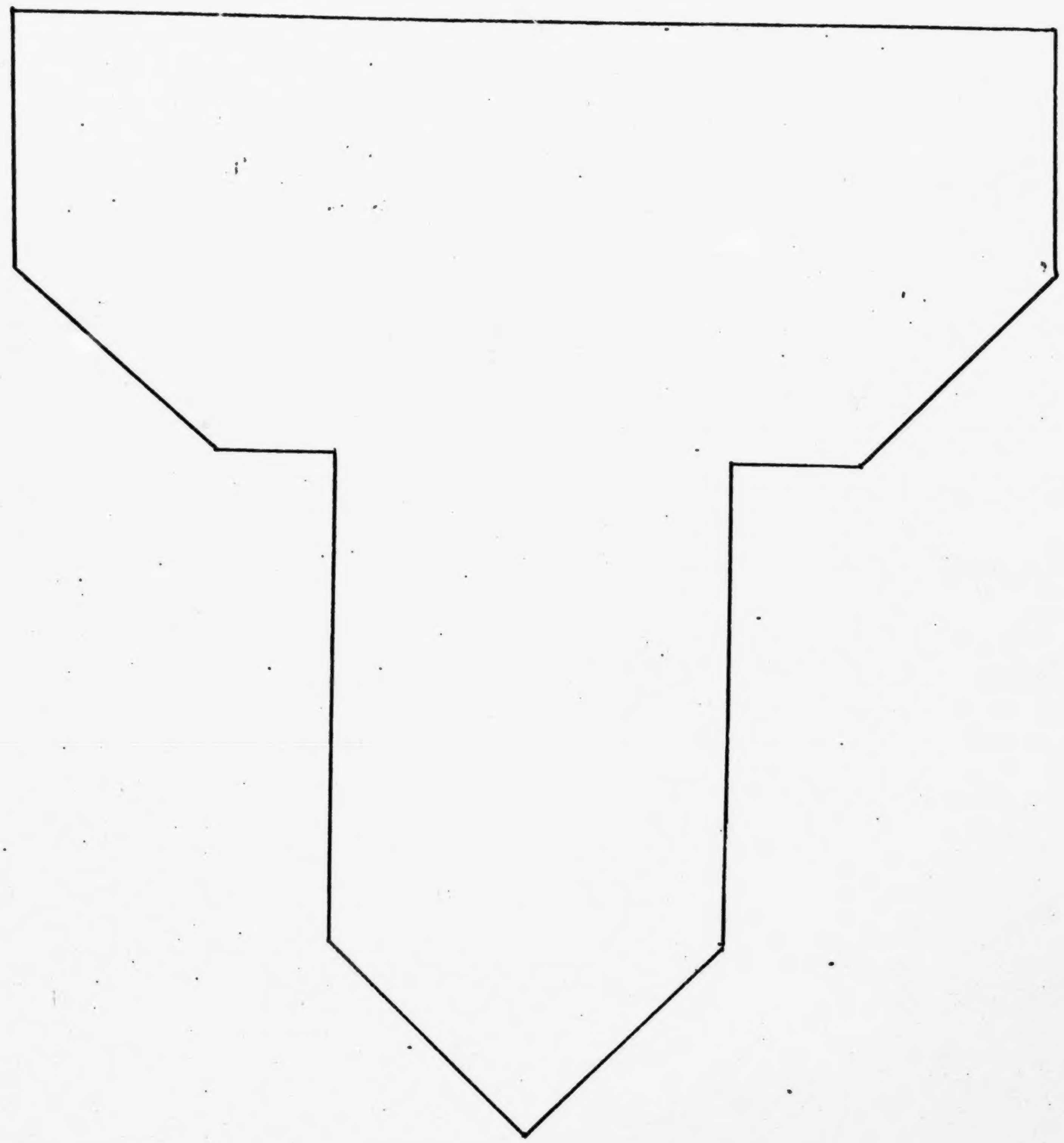


Figure 3

Solution for Design (B)--Task (2)

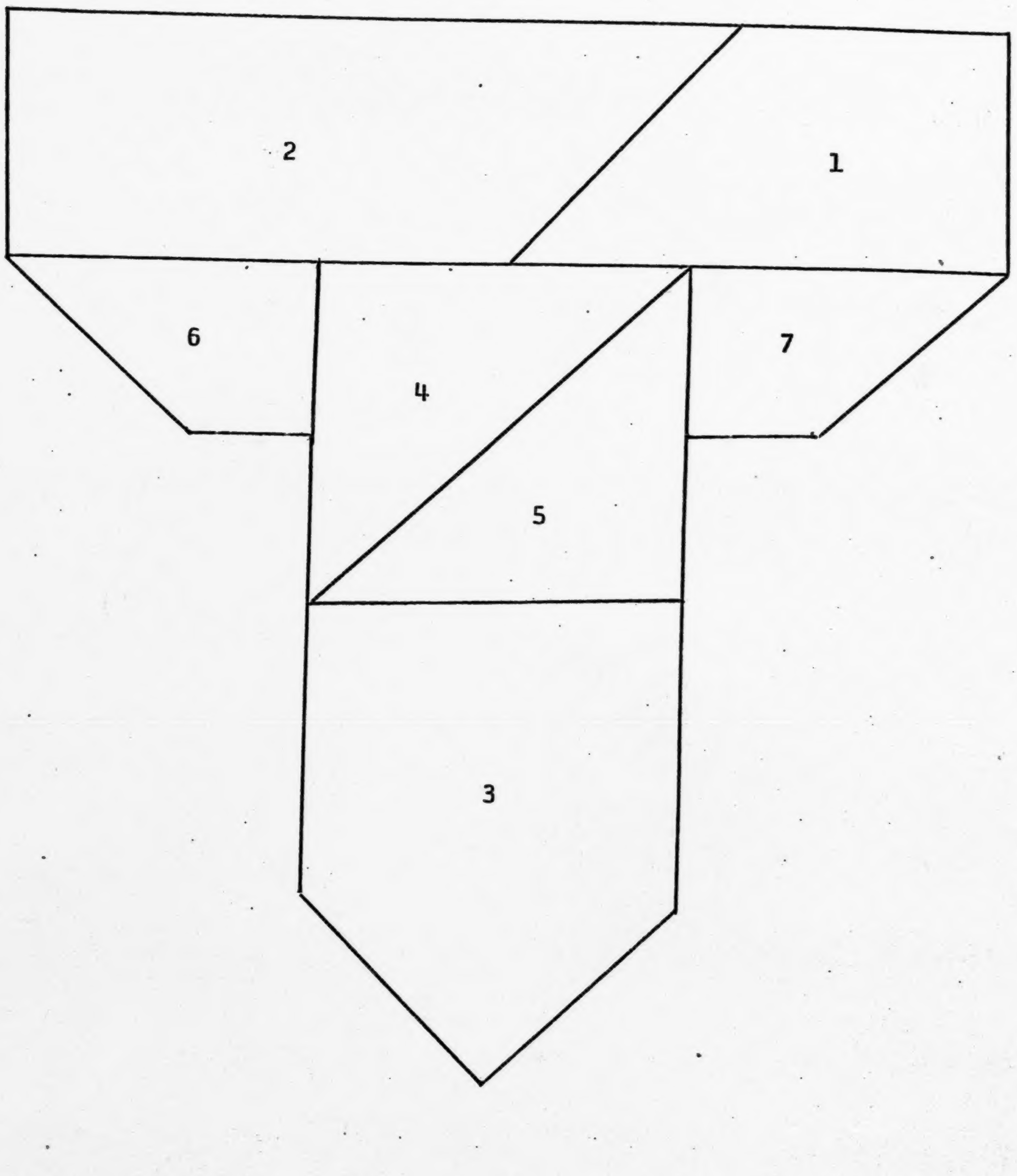


Figure 4

Partial Reproduction of Card in Set (A), Task (3)

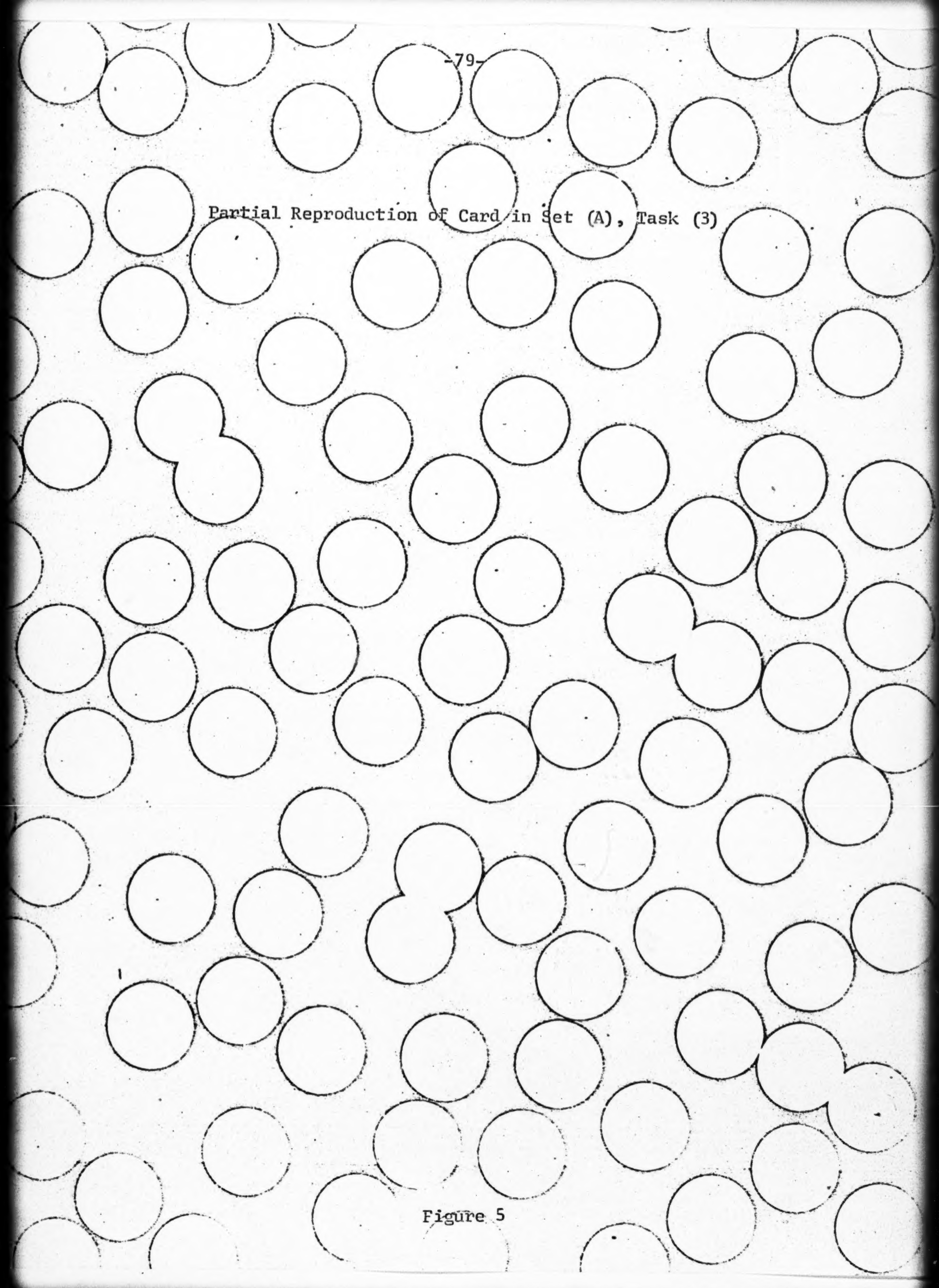


Figure 5

Partial Reproduction of Card in Set (B), Task (3)

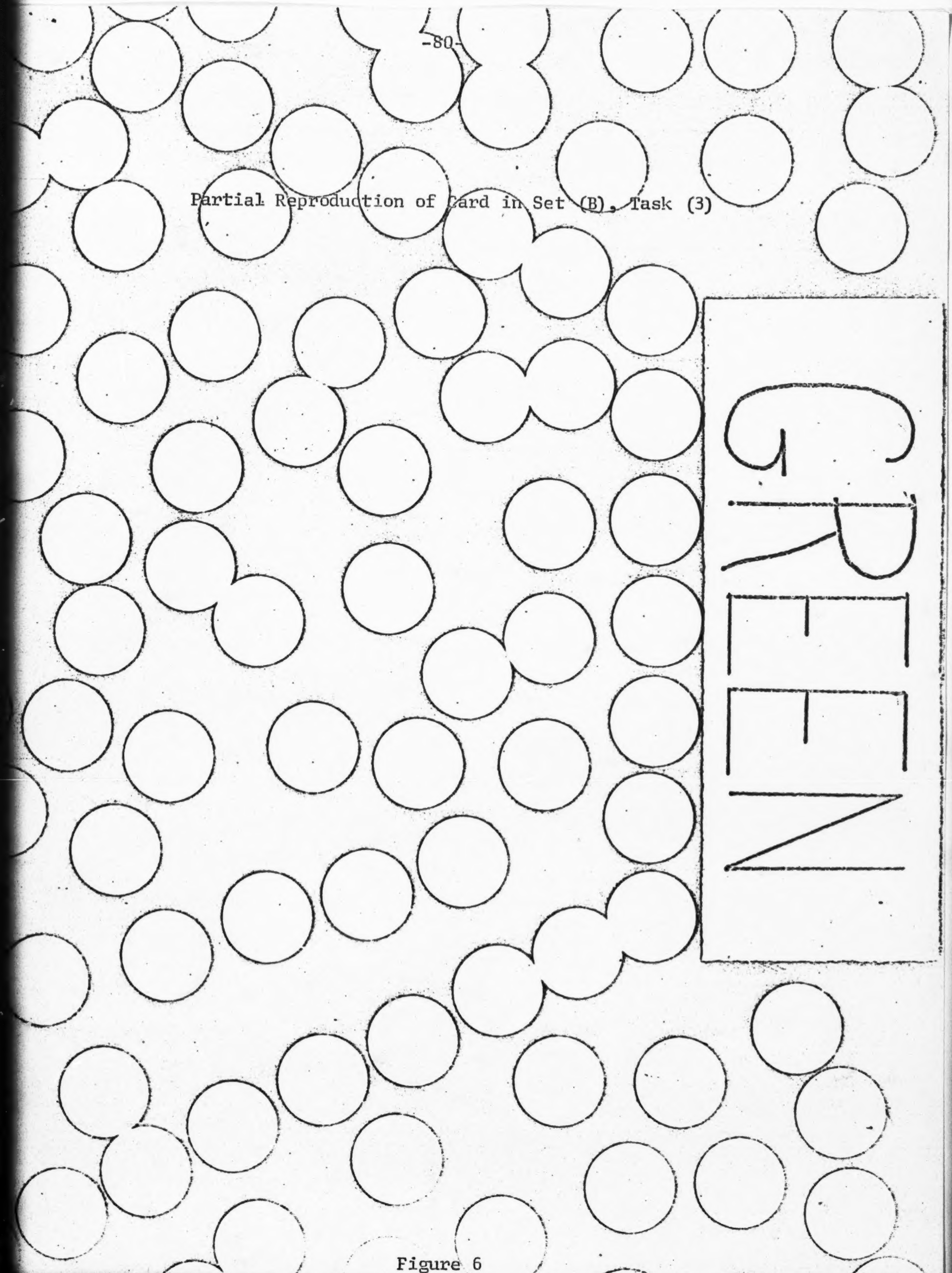


Figure 6

Partial Reproduction of Card in Set (C), Task (3)

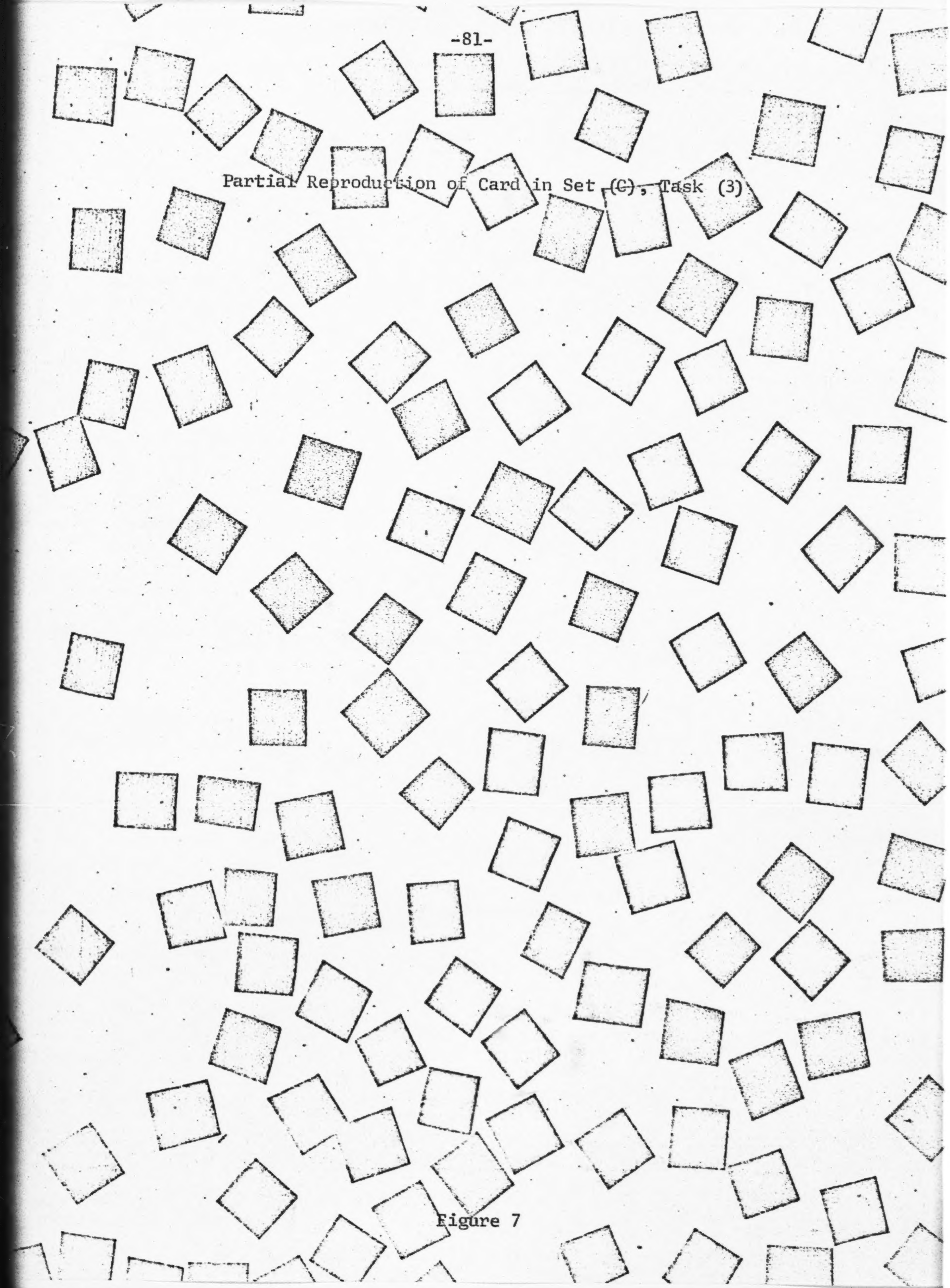
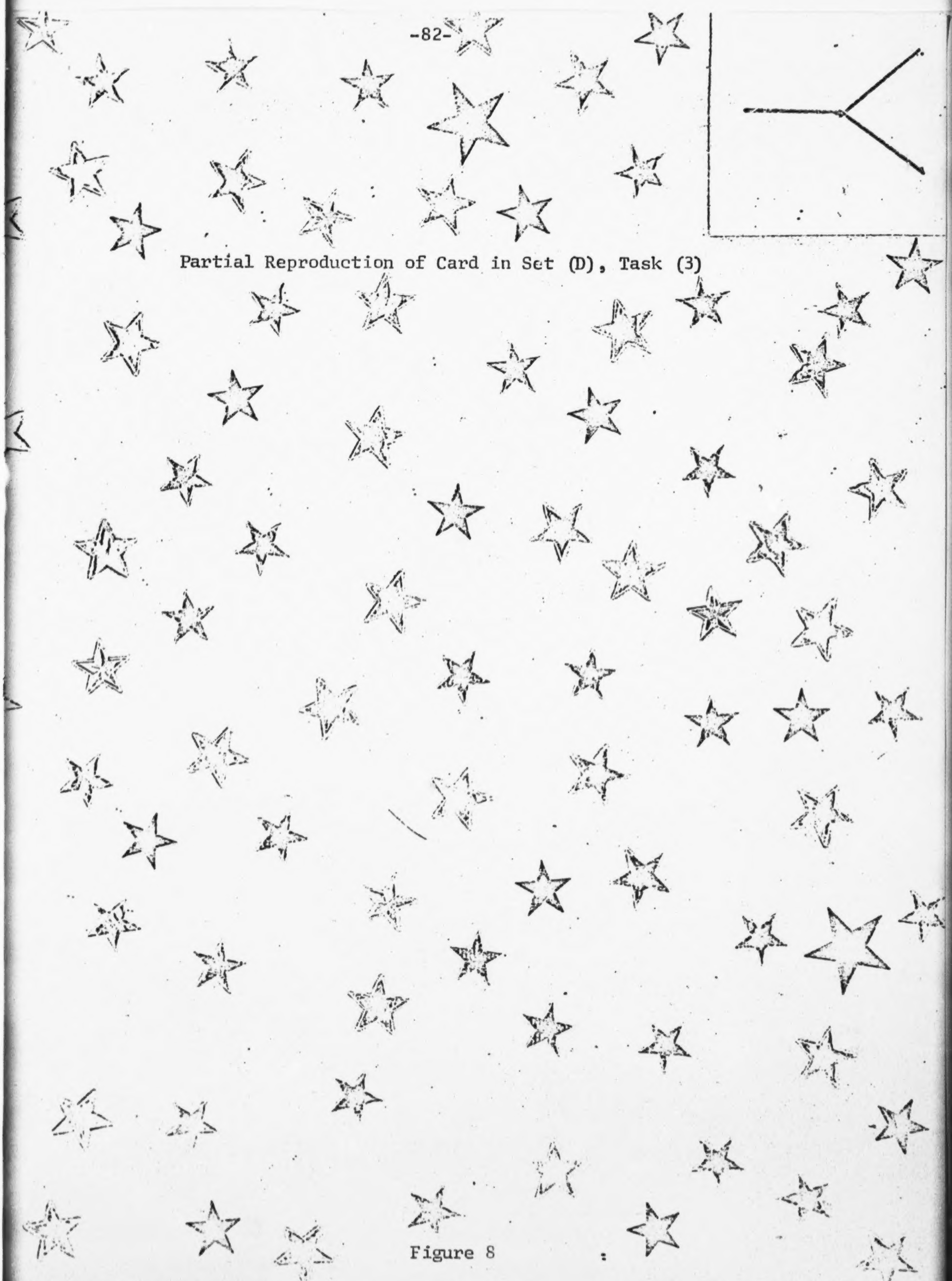


Figure 7



Partial Reproduction of Card in Set (D), Task (3)

Figure 8

Cup Stimulus, Set (E), Task (3)

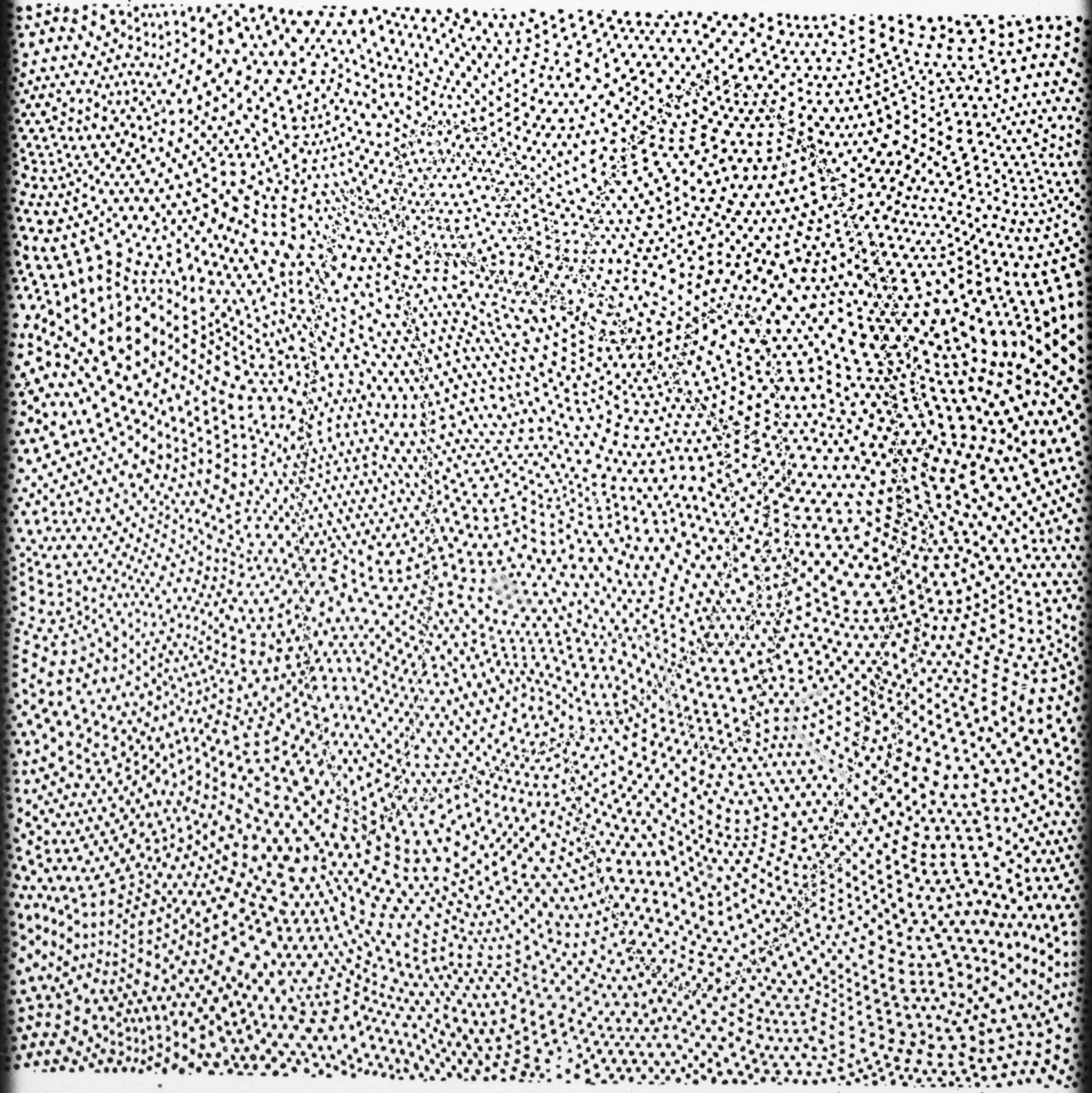


Figure 9

Horse Stimulus, Set (F), Task (3)

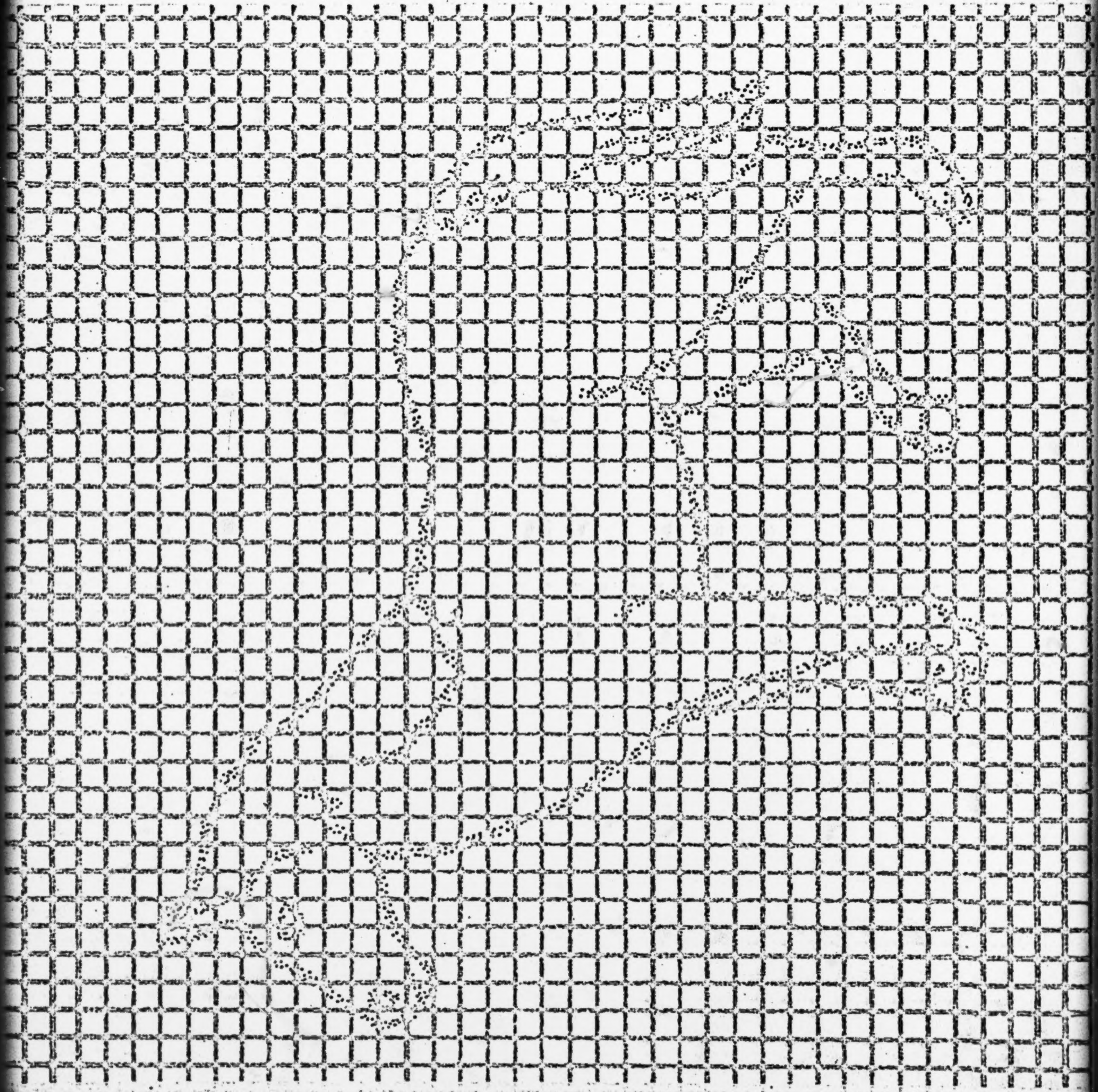
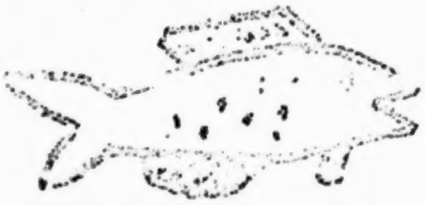
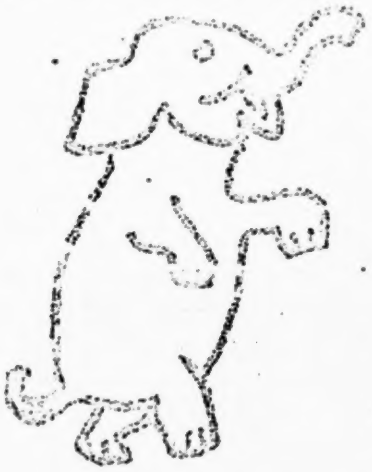
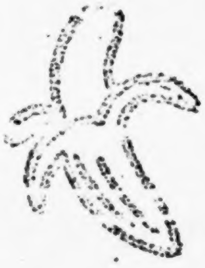


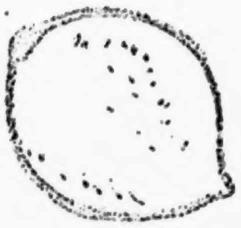
Figure 10

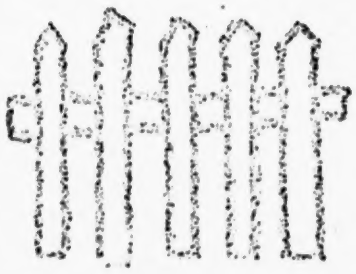
A Page from Booklet, Task (4)

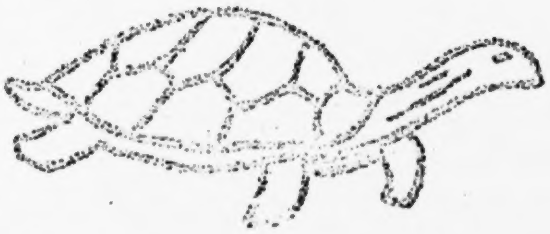


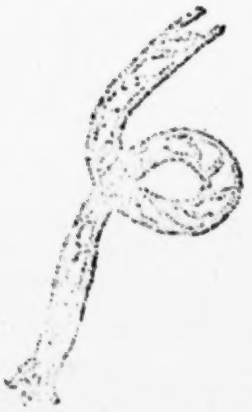


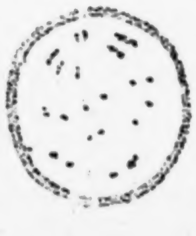












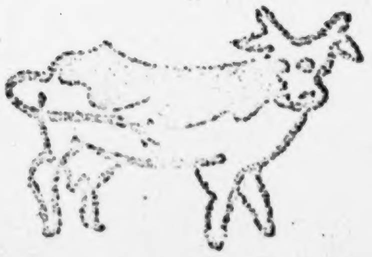


Figure 11

Doll-Like Figure used for Demonstration Purposes, Task (5)

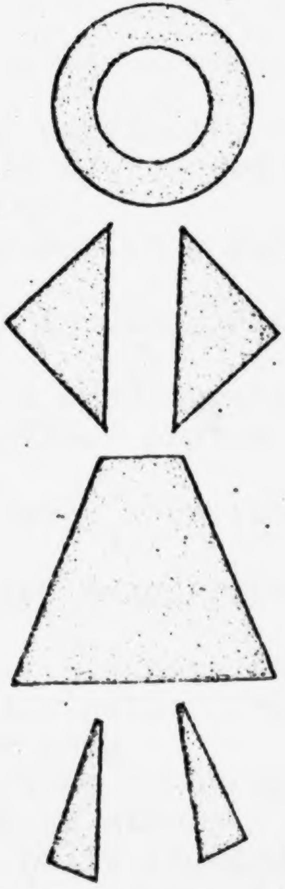


Figure 12

Appendix A

The Classroom Behavior Checklist

CODE __________ SEX __________ DATE __________

OBSERVER __________ TIME STARTED __________

OBSERVATION NUMBER __________

SCHOOL __________ TEACHER ____________________

	Frequently	Moderately Often	Rarely Often	Never	No Opportunity to Observe; Can't Rate
(1) withdraws in the face of difficulty					
(2) is inattentive; is easily distracted by things going on around him					
(3) uses materials and equipment in a rough or destructive manner					
(4) volunteers contributions to class discussions and projects					
(5) helps, sympathizes, shows consideration of others; is thoughtful; offers praise, guidance, assurance					
(6) sits quietly and attentively when task requires					
(7) brags, shows off, displays exaggerated opinion of own abilities					
(8) seeks attention, i.e., in a manner that seems designed for active notice rather than stemming from a need for help					
(9) utilizes available free time constructively					
(10) interferes with the work of others					
(11) withstands interference while engaged in individual activities, e.g., doing difficult assignment, a puzzle, painting					
(12) imitates others rather than initiating own activities					
(13) gets upset when called upon in class; appears fearful, tense, timid					
(14) takes appropriate initiative in problem-solving situations					
(15) seeks to be near others					
(16) does not follow directions					
(17) requires close assistance or direction in order to work at a task					
(18) seeks to do things differently from others, even when own method is not effective					
(19) lets other children impose on him or boss him around					
(20) tries to figure out things for himself before calling on teacher (leader) or other children for help					

(21) enters readily into new situations; is venturesome, inquisitive, etc., when appropriate to ongoing activities
(22) does not pursue task to completion
(23) seeks constant recognition and/or approval
(24) speaks out of turn; interrupts others; takes other's turn, etc.
(25) attempts to resolve difficulties that arise between himself and other children without appealing to teacher (leader)
(26) approaches new tasks timidly and without confidence; shrinks from trying new things
(27) tends to choose habitual and familiar activities

Appendix B

The Revised Classroom Behavior Checklist

CODE ______ SEX ______ OBSERVER ______ DATE ______ TIME STARTED ______ OBSERVATION NUMBER ______

SCHOOL ______ TEACHER ______

		Frequently	Moderately Often	Rarely	Never	No Opportunity
(1)	withdraws in the face of difficulty					
(2)	is inattentive; is easily distracted by things going on around him					
(3)	uses materials and equipment in a rough or destructive manner					
(4)	helps, sympathizes, shows consideration of others; is thoughtful; offers praise, guidance, assurance					
(5)	sits quietly and attentively when task requires					
(6)	brags, shows off, displays exaggerated opinion of own abilities					
(7)	seeks attention, i.e., in a manner that seems designed for active notice rather than stemming from a need for help					
(8)	utilizes available free time constructively					
(9)	interferes with the work of others					
(10)	withstands interference while engaged in individual activities, e.g., doing difficult assignment, a puzzle, painting					
(11)	does not follow directions					
(12)	requires close assistance or direction in order to work at a task					
(13)	seeks to do things differently from others, even when own method is not effective					
(14)	lets other children impose on him or boss him around					
(15)	tries to figure out things for himself before calling on teacher (leader) or other children for help					
(16)	does not pursue task to completion					
(17)	seeks constant recognition and/or approval					
(18)	speaks out of turn; interrupts others; takes other's turn, etc.					
(19)	attempts to resolve difficulties that arise between himself and other children without appealing to teacher (leader)					